Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *University of Michigan*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Success Probability and Choice Behavior: N. T. FEATHER 257

Stimulus Correlations for the Judged Illumination of a Surface: J. BECK 267

Time Estimates as Affected by Motivational Level, Goal Distance, and Rate of Progress: R. D. MEADE 275

Inhibitory Set in Problem Solving as Related to Reinforcement Learning: R. ADAMSON 280

Amount of Conditioning and Intensity of the Conditioned Stimulus: H. D. KIMMEL 283

Stimulus Similarity and Task Familiarity as Determinants of Expectancy Generalization: D. HEATH 289

Frustration Stimuli in Discrimination: D. W. TYLER, M. H. MARX, AND G. COLLIER ... 295

The Effect of "Social" Discriminative Cues on Probability Learning: E. D. NEIMARK AND S. ROSENBERG 302

The Effect of Shock on the Rat's Choice of a Path to Food: T. F. LOHR 312

Supplementary Report: Effect of Interpolated UCS Trials in Eyelid Conditioning Without a Ready Signal: K. P. GOODRICH, L. E. ROSS, AND A. R. WAGNER 319

American Psychological Association
Vol. 58, No. 4 October 1959

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*.** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1959 the publication cost is $30 per printed page. Authors of priority publications receive no free offprints.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1959 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 58, NO. 4 OCTOBER, 1959

SUCCESS PROBABILITY AND CHOICE BEHAVIOR

N. T. FEATHER [1]

University of New England, New South Wales, Australia

The present investigation is concerned with the effect of varying a person's subjective probability of attaining a goal in a choice situation involving different goal objects. Specifically, three problems which appear to be basic ones in the analysis of choice behavior are investigated. They are: (*a*) Is the attractiveness of attaining a goal object related to *S*'s subjective probability of attaining it? (*b*) Is the tendency to choose to get a goal object related to *S*'s subjective probability of attaining it? (*c*) Might one expect these relationships to vary between, on the one hand, an *ego-related* situation (one in which success can obviously be related to the efforts and skill of the person) and a *chance-related* situation (one in which success is apparently determined by chance factors),[2] and on the other hand, between an *achievement-oriented* situation (a test-like situation involving considerable pressure towards doing well) and a *relaxed* situation (a freer, more play-like situation involving no obvious pressure towards high achievement)?[3] In this investigation an attempt is made to elucidate the above relationships and to examine whether they vary for the types of situation indicated.

Three concepts are employed in this study: (*a*) *Attainment attractiveness:* Given a number of goal objects A, B, C . . . the attainment attractiveness of A> attainment attractiveness of B> . . ., if in a situation in which *S* is required to express his wishes with respect to attainment, and in the absence of any commitment to the choice, he states that he would prefer to get A, rather than to get B . . . (*b*) *Choice potential:* Given a number of goal objects A, B, C, . . . the choice potential associated with A> choice potential associated with B> . . . if, in a situation in which there is the implication that *S* is committed to his choice, he chooses

[1] Now on leave at the University of Michigan, Ann Arbor, Michigan. The author thanks Duncan Howie for his advice in connection with this research which was submitted in fulfillment of the Honor's M.A. degree at the University of New England, New South Wales.

[2] This parallels Filer's (1952) distinction between merit and nonmerit conditions in his study of factors affecting the attractiveness of goal objects and a recent distinction made by Phares (1957) between skill and chance situations.

[3] This parallels the distinction made by McClelland et al. (1953) and Atkinson (1954) in their investigations of the achievement motive.

to get A, rather than B, . . . (*c*) *Success probability:* The success probability associated with a goal object refers to the probability held by *S* that a particular goal object may be obtained.

Attainment attractiveness is analogous to Lewin's (1938, p. 88) valence, to Rotter's (1954, p. 107) reinforcement value, and to the concept of utility in decision theory (Edwards, 1954). Choice potential may be related to Lewin's (1938, p. 89) force, to Rotter's (1954, p. 105) behavior potential, to subjectively expected utility (SEU) in decision theory (Edwards: 1954, 1955) and to Atkinson's (1957) strength of motivation. It will be noted that the definitions of both attainment attractiveness and choice potential refer to the *attainment* of the goal object rather than to the goal object per se, and the operational distinction between these concepts is in terms of a wishful choice as opposed to a choice which may have to be acted upon. Success probability parallels Rotter's (1954, p. 107) expectancy, the expectancies of Tolman (1955) and Atkinson (1957), and also corresponds to the concept of subjective probability employed in Lewinian theory (Lewin, 1943; Lewin, Dembo, Festinger, & Sears, 1944) and in decision models (Edwards, 1954; Coombs & Beardslee, 1954).

In terms of these concepts two hypotheses are advanced. *Hypothesis 1* is that the attainment attractiveness of a goal object should tend to vary inversely with the associated success probability. This assumed covariation should tend to be more apparent in ego-related than in chance-related situations, and more apparent in achievement-oriented than in relaxed situations. *Hypothesis 2* is that the choice potential associated with a goal object should tend to vary directly with the associated success probability. This assumed covariation should tend to be less apparent in ego-related than in chance-related situations, and less apparent in achievement-oriented than in relaxed situations. Thus, in Hypothesis 1 it is predicted that in a situation free from commitment, a person should tend to wish to get a goal object more as success becomes less likely for him. In Hypothesis 2 it is predicted that in a situation implying commitment, a person should be less inclined to choose to get a goal object as success becomes less likely for him. These relationships are, in turn, considered in terms of situational context.

Assumptions involved in these hypotheses are presented in detail elsewhere (Feather, 1959). For Hypothesis 1 these assumptions consider the increase in the value placed on achievement with increase in difficulty level, and the relationship of this achievement value to different situations, the assumption being that success is valued more in an ego-related situation than in a chance-related situation, and more in an achievement-oriented situation than in a relaxed situation. For Hypothesis 2 it is assumed that past experiences in commitment situations where there is the possibility of loss may lead to a tendency to choose the easy rather than the difficult, i.e., success probability should tend to have a *constraining* effect on choice. However, the value placed upon the difficult achievement is also assumed to be an important factor in a committed choice. Thus it may happen that this value becomes the focal element in decision in ego-related and achievement-oriented situations, even though the choice involves commitment. In short, it may be found

that in these situations, committed choice tends to follow *S*'s wishes.

METHOD

Subjects and Design

The variables were incorporated into a $2 \times 2 \times 2$ factorial design giving eight groups with treatments listed in Table 1.

Twenty-four boys, sampled from Grade 6 of New South Wales primary schools, were randomly assigned to each group. It was considered that at the age level sampled, one might reasonably expect to find evidence of the effect of learning on choice behavior as indicated in the assumptions. Further, sampling from an exclusively male population controls the possibility of sex differences, which might arise due to differences in the upbringing of girls compared with boys. Analyses of variance indicated that differences between groups with respect to ages and IQs are not significant.

General Procedure

Five candies, selected on the basis of pilot research to be similar in attractiveness, were used as goal objects. They were of the Life-saver type (lime, orange, butterscotch, barley sugar, lemon). The presentation was by way of paired-comparisons, each *S* being presented with the 10 pairs of candies successively in an individual testing situation. One candy in each pair was easier to get than the other, the difference in success probability between the candies being related to different task requirements T_1 and T_2. The basic order used in presenting the 10 pairs was derived from Ross' (1934) lists to meet the usual requirements (Guilford, 1954, p. 160) of a paired-comparison arrangement. These pairs were also arranged so that the comparisons into which a particular candy entered when associated with T_1 were the same as the comparisons involved when this candy was associated with T_2.

The *S* was required to make his selection from each pair in accordance with the dependent variable studied. For attainment attractiveness he was asked, "Now if you were to get one of these, which one would you feel most pleased about getting? Which one do you *wish* you could get the most?" For choice potential the question was, "Which one would you actually *choose* to try to get?"

The *S* was asked to give a reason for his selection from each pair immediately after each choice, thus providing a set of qualitative data for analysis.

TABLE 1

TREATMENTS FOR THE EIGHT GROUPS

Dependent Variable		Independent Variable: Situation
Attainment Attractiveness	Choice Potential	
Group A	Group E	Ego-related +Achievement-oriented
Group B	Group F	Chance-related +Achievement-oriented
Group C	Group G	Ego-related+Relaxed
Group D	Group H	Chance-related +Relaxed

Varying Success Probability

The operational specification of success probability involved making attainment contingent upon satisfying a certain performance requirement in a task, and variations in success probability were achieved by selected changes in this task requirement.

Ego-related situation.—Under ego-related conditions the task used was card sorting. The *Ss* were required to sort a pack of 36 playing cards into three trays, face cards and 6's being excluded. It was assumed that since this task obviously involves learning, and improvement is gradual, *S* should tend to see his results as largely due to his own efforts and skill. The *S* worked at this task for six trials for which fictitious prearranged time scores were reported. A realism was given to these reported scores by basing them on the trend of actual scores obtained in a small pilot study. It was thought that the use of a novel task (card sorting), in conjunction with a prearranged sequence of scores, would allow sharper control over success probability. Two sequences of scores (Sequence I: 57, 54, 51, 49, 48, 47; and Sequence II: 58, 55, 52, 50, 49, 48) were used so as to reduce the risk that posttest comparisons of scores by *Ss* might lead to the detection of their fictitious nature. Each of these sequences was used for half the *Ss*.

Following the six trials on the card-sorting task the 10 pairs of candies were introduced in succession. In each pair, one candy was associated with a task requirement T_1 (41 sec. following Sequence I; 42 sec. following Sequence II); the other candy with a task requirement T_2 (46 sec. following Sequence I; 47 sec. following Sequence II). These task

requirements were selected on the basis of pilot research so as to determine differences in success probability. The *S* was led to believe that attainment of a candy depended upon fulfilling the associated performance requirement. Thus, "Suppose you were actually going to try to get one of these, and that you would get it only if you could sort the cards within a certain time. If you decided to try for this one, you would get it only if you sorted the cards within 41 sec. (T_1). If you tried for this one, you would get it only if you sorted the cards within 46 sec. (T_2). So this packet is an *easy* one to get, this one is a *hard* one to get."

Chance-related situation.—Here the task used was a lottery box containing 15 red marbles and 5 blue marbles. One of the marbles could be rolled through a gate at the bottom of the box. For each of the 10 pairs of candies presented, one candy was associated with a task requirement T_1 (rolling out a blue marble), the other candy with a task requirement T_2 (rolling out a red marble). Each *S* was led to believe that attainment of a candy depended upon rolling out a marble of a particular color. Thus, "If you decided to try for this one you would get it only if you rolled a blue marble out of the box (T_1). If you tried for this one you would get it only if you rolled a red marble out of the box (T_2)."

It will be observed that for both ego-related and chance-related situations, in each pair, the candy associated with T_1 was more difficult to get than the candy associated with T_2.

Achievement-oriented and Relaxed Situations

Under achievement-oriented conditions card sorting was introduced as a test of speed and the lottery box as a test of luck. In the relaxed situation, card sorting was introduced as a *game* played with cards and the lottery box as a marble game.

Statistical Analysis

For each group the frequency of selection (+) and nonselection (−) of each candy when associated with task requirements T_1 and T_2 was tabulated. Sutcliffe's (1957) extension of the χ^2 technique to frequency data for multiple classification designs, was applied to the data to bring out interaction effects.[4]

[4] The author is indebted to J. P. Sutcliffe for his advice in the application of this analysis. In terms of his discussion the present design is of the 2(b) form, the mixed model with parameters estimated from the data.

Investigation of Actual Choice

In the investigation of choice potential in the above experiment, the *S* was not, in fact, required to act upon his choice, although a commitment type of question was used. The relationship between choice potential and success probability was also investigated in a situation of *actual* choice. One might expect that, under these conditions, the constraining effect of success probability on choice would be even more pronounced.

Thus *E* also presented each *S* with a pair of candies, one associated with the T_1 task requirement, the other with the T_2 task requirement in an actual choice situation. The same pair of candies was used throughout for all groups. Each *S* was taken individually not more than one week after the main session and was presented with these candies arranged in the same way as in the main session. The same conditions with respect to ego-related or chance-related, achievement-oriented or relaxed situations, as in the main session, also applied. Instructions emphasized an actual choice: "I'm going to let you actually try to get one of these today. Which one are you *actually* going to try to get?"

Results

Analysis of Frequency Data

The classifications are symbolized as follows: *E* (ego-related situation), *C* (chance-related situation); *A* (achievement-oriented situation), *R* (relaxed situation); T_1 (difficult task requirement), T_2 (easy task requirement); + (selected), − (not selected).

Because of the pairing system employed, when one candy is selected (+) another is rejected (−). Hence, the five candies are not included as another classification in the χ^2 analysis of results since they are not independent with respect to the type of tabulation employed. In view of this, the analysis of results bearing on Hypotheses 1 and 2 is made for each candy separately, and similarities/dissimilarities in the five different analyses are indicated.

Hypothesis 1.—Results from Groups A, B, C, and D are relevant to Hypothesis 1. Detailed presentation is made in Table 2 of the results for Candy D (barley sugar).

Table 3 summarizes the results of the χ^2 analysis applied to this data. The χ^2 values which are particularly relevant to Hypothesis 1 are those from Sources 3, 5, 6, and 7 in Table 3. It will be noted that all but the last are highly significant.

In Table 4 there is a significant *decrease* in the proportion of + selections (i.e., less of a tendency to wish to get Candy D), as the task requirement changes from T_1 to T_2 (i.e., an *increase* in the success probability associated with Candy D). Thus, for conditions considered as a whole, there is evidence that the attainment attractiveness of Candy D is inversely related to success probability. This supports Hypothesis 1. Table 4 also shows that the decrease in the proportion of + selections from T_1 to T_2 under ego-related conditions is greater than the decrease under chance-related conditions. Similarly, this decrease is seen to be greater under achievement-oriented conditions than under relaxed conditions. The results show, then, that the inverse relationship

TABLE 2

FREQUENCIES OF SELECTION (+) AND NONSELECTION (−) OF CANDY D FOR ALL CONDITIONS

Condition	Choices		Total
	Selected (+)	Not Selected (−)	
Ego, Achievement, T_2	1	47	48
Ego, Achievement, T_1	40	8	48
Ego, Relaxed, T_2	11	37	48
Ego, Relaxed, T_1	30	18	48
Chance, Achievement, T_2	12	36	48
Chance, Achievement, T_1	33	15	48
Chance, Relaxed, T_2	24	24	48
Chance, Relaxed, T_1	26	22	48
Totals	177	207	384

Note.—Table 2 is to be read as follows: In the ego-related, achievement-oriented situation Candy D is selected once and rejected 47 times in the 48 pairs in which its attainment depends upon satisfying the task requirement T_2 (sorting the cards within 46 or 47 sec.) and when attainment attractiveness is the dependent variable.

TABLE 3

RESULTS OF χ^2 ANALYSIS FOR CANDY D

Source	χ^2	*df*	*P*
$(EC)(+-)$	1.77	1	>.05
$(AR)(+-)$	.26	1	>.05
$(T_1T_2)(+-)$	68.76	1	<.001
$(EC)(AR)(+-)$	.26	1	>.05
$(EC)(T_1T_2)(+-)$	12.84	1	<.001
$(AR)(T_1T_2)(+-)$	15.94	1	<.001
$(EC)(AR)(T_1T_2)(+-)$	.01	1	>.05
Total	99.84	7	<.001

TABLE 4

FREQUENCIES OF SELECTION (+) AND NONSELECTION (−) OF CANDY D RELATING TO $\chi^2_{(T_1T_2)(+-)}$, $\chi^2_{(EC)(T_1T_2)(+-)}$, AND $\chi^2_{(AR)(T_1T_2)(+-)}$

Condition	Choices		Total
	Selected (+)	Not Selected (−)	
T_2	48	144	192
T_1	129	63	192
Total	177	207	384
Ego, T_2	12	84	96
Ego, T_1	70	26	96
Chance, T_2	36	60	96
Chance, T_1	59	37	96
Total	177	207	384
Achievement, T_2	13	83	96
Achievement, T_1	73	23	96
Relaxed, T_2	35	61	96
Relaxed, T_1	56	40	96
Total	177	207	384

TABLE 5

FREQUENCIES OF SELECTION (+) AND NONSELECTION (−) OF CANDY A FOR ALL EXPERIMENTAL CONDITIONS

Condition	Choices		Total
	Selected (+)	Not Selected (−)	
Ego, Achievement, T_2	17	31	48
Ego, Achievement, T_1	33	15	48
Ego, Relaxed, T_2	41	7	48
Ego, Relaxed, T_1	12	36	48
Chance, Achievement, T_2	20	28	48
Chance, Achievement, T_1	26	22	48
Chance, Relaxed, T_2	35	13	48
Chance, Relaxed, T_1	9	39	48
Totals	193	191	384

Note.—Table 5 is to be read as follows: In the ego-related, achievement-oriented situation Candy A is selected 17 times and rejected 31 times in the 48 pairs in which its attainment depends upon satisfying task requirement T_2 (sorting the cards within 46 or 47 sec.) and when choice potential is the dependent variable.

between the attainment attractiveness of Candy D and success probability is significantly more apparent under ego-related and achievement-oriented conditions. Again this supports Hypothesis 1. Finally, in Table 2 there is no significant variation in the $\chi^2_{(AR)(T_1T_2)(+-)}$ interaction picture from ego-related to chance-related conditions.

The analyses for each of the other four candies reveal the same pattern of χ^2 values as that reported for Candy D.

Hypothesis 2.—Results from Groups E, F, G, and H are relevant to Hypothesis 2. Detailed presentation is made in Table 5 of the results for Candy A (lime).

The χ^2 analysis applied to this data shows that: $\chi^2_{(T_1T_2)(+-)} = 11.34$ ($df = 1$, $p < .001$), $\chi^2_{(EC)(T_1T_2)(+-)} = .51$ ($df = 1, p > .05$), $\chi^2_{(AR)(T_{12}T)(+-)} = 61.76$ ($df = 1$, $p < .001$), and $\chi^2_{(EC)(AR)(T_1T_2)(+-)} = 1.76$ ($df = 1$, $p > .05$).

In Table 6 there is a significant *increase* in the proportion of + selections (i.e., more of a tendency to choose to get Candy A), as the task requirement changes from T_1 to T_2 (i.e., an *increase* in the success probability associated with Candy A). Thus, for conditions considered as a whole, there is evidence that the choice potential associated with Candy A is directly related to success probability. This supports Hypothesis 2. However, in Table 6, there is no significant variation in this $\chi^2_{(T_1T_2)(+-)}$ interaction picture over the levels of the ego-related, chance-related classification, and hence no evidence that the direct relationship between the choice potential associated with Candy A and success probability is significantly less apparent under ego-related

TABLE 6

FREQUENCIES OF SELECTION (+) AND NONSELECTION (−) OF CANDY A RELATING TO $\chi^2_{(T_1T_2)(+-)}$, $\chi^2_{(EC)(T_1T_2)(+-)}$ AND $\chi^2_{(AR)(T_1T_2)(+-)}$

Condition	Choices		Total
	Selected (+)	Not Selected (−)	
T_2	113	79	192
T_1	80	112	192
Total	193	191	384
Ego, T_2	58	38	96
Ego, T_1	45	51	96
Chance, T_2	55	41	96
Chance, T_1	35	61	96
Total	193	191	384
Achievement, T_2	37	59	96
Achievement, T_1	59	37	96
Relaxed, T_2	76	20	96
Relaxed, T_1	21	75	96
Total	193	191	384

conditions. This part of Hypothesis 2 is therefore not supported by the results. Table 6 also shows that, while there is an increase in the proportion of + selections from T_1 to T_2 under relaxed conditions, the reverse happens under achievement-oriented conditions. Thus, under achievement-oriented conditions for Candy A, the relationship between choice potential and success probability is no longer direct, and, in fact, becomes *inverse*. Hence, there is very definite support for that part of Hypothesis 2 which states that the direct relationship should tend to be less apparent in the achievement-oriented situation. Finally, in Table 5 there is no significant variation in the $\chi^2_{(AR)(T_1T_2)(+-)}$ interaction picture from ego-related to chance-related conditions.

The analyses for each of the other four sweets reveal exactly the same pattern of χ^2 values with some inconsistency for $\chi^2_{(EC)(AR)(T_1T_2)(+-)}$, this value being significant for one of the candies ($p < .05$), and approaching significance ($p < .10$) for two others.

Analysis of Reasons

Reasons were classified for the different experimental conditions in terms of four categories: P (the reason referred to probability of attainment), R (the reason referred to the attractiveness of the candy), A (the reason referred to achievement factors) and M (a mixture of the above). Sutcliffe's (1957) extension of the χ^2 analysis was again applied to the data, separate analyses being made for reasons bearing on Hypothesis 1 and reasons bearing on Hypothesis 2.[5] The results are in general agreement with the assumptions involved in the hypotheses. For example, wishing to get the difficult is most often explained in terms of "wanting to do well" (an A type of reason), while choosing the easy is generally explained in terms of the probability of success, "the easy one to get" (a P type of reason).

[5] Complete tables of frequencies of the four categories of reasons for each condition in the experiment have been deposited with the American Documentation Institute. Order document No. 6025, submitting $1.25 for photo prints or $1.25 for microfilm.

TABLE 7

CHANGES IN SELECTION FROM VERBAL TO ACTUAL SITUATIONS

Group	T_1		T_2		P
	T_2	T_1	T_2	T_1	
A	7	15	2	0	.008
B	6	11	7	0	.016
C	9	8	7	0	.002
D	8	2	14	0	.004
E	8	7	8	1	.020
F	6	9	9	0	.016
G	3	2	17	2	.500
H	1	4	18	1	>.188

Note—Table 7 is to be read as follows: For Group A, 7 *S*s who selected the T_1 or less likely candy on the first or verbal occasion changed to choice of the T_2 or more likely candy on the second or actual occasion.

Analysis of Actual Choice

For the pair of candies used, one can compare *S*'s selection in the first or *verbal* situation with his choice in the second or *actual* situation, other conditions remaining the same. For each group the significance of change from one situation to the other is tested by the binomial expansion. Since it is predicted that changes in selection under actual conditions should tend to be in the direction of choice of the T_2 or easier alternative, a one-tailed test is used, and the significance values are taken from Siegel (1956, p. 250).

The results are summarized in Table 7.

Table 7 shows that all changes are significant for Groups A, B, C, and D. For these groups the over-all

change is from a situation involving a statement of wishes concerning attainment, to one in which *S* is actually committed to his choice. Under these circumstances there is a tendency to change to the choice of the T_2 or more likely candy in the actual situation, and this is in line with expectation.

For Groups E, F, G, and H, choice potential is the dependent variable in both verbal and actual situations. Hence, for each of these groups, the changes in selection may be directly related to the verbal-actual difference. The changes are significant for Groups E and F, but are not significant for Groups G and H. Groups E and F are similar in that each is associated with achievement-oriented conditions. It would seem that, for these groups, there is evidence for a greater concern about the probability of attainment when the choice becomes more actual. Choice is more in the direction of the easier alternative and this is in line with expectation. As against this, when relaxed conditions are common, as for Groups G and H, it would seem that *S*'s choice in the actual situation is likely to correspond to his choice in the verbal situation, the predominant tendency under both conditions being to choose the easier alternative.

Discussion

Within the scope of this investigation, these results generally support Hypotheses 1 and 2. There is evidence for a greater tendency to wish to get a goal object as its attainment becomes subjectively *less* likely in a situation where commitment is absent, and a greater tendency to choose to get a goal object as its attainment becomes subjectively *more* likely in a situation where commitment to the choice is implied. These relationships are also seen to be related to situational context. The reasons given by *S*s for their selections indicate the importance for them of goal object attractiveness, achievement values, and the constraint of success probability as factors affecting their decisions in the different situations. Finally, the constraining effect of success probability on choice behavior under commitment conditions becomes particularly apparent in a situation where, in fact, *S* has to act upon his choice.

The results of this study may be related to five theoretical discussions in the literature: (*a*) the analysis of level of aspiration behavior by Lewin et al. (1944), (*b*) Tolman's (1955) discussion of the principles of performance, (*c*) Rotter's (1954) discussion of the relationships between his basic concepts of behavior potential, reinforcement value and expectancy, in his social learning theory, (*d*) decision models by Edwards (1955) and by Coombs and Beardslee (1954) which involve the concepts of utility and subjective probability, and (*e*) Atkinson's (1957) analysis of risk-taking behavior. In these analyses from different areas of research, we find that similar types of concept are used and that they are related in similar types of equation.

The inverse relationship of Hypothesis 1 is consistent with the inverse relationship between positive valence of successful attainment and subjective probability in the analysis of level of aspiration behavior by Lewin et al. (1944, pp. 360–361). It is also consistent with Atkinson's (1957) inverse relationship between incentive value of achievement and subjective probability. It is, however, contrary to Rotter's (1954, p. 107) assumption that reinforcement value and expectancy are independent, and to the similar assumption implied by Edwards (1954, p. 400) and made by Coombs and Beardslee (1954, p. 263), that values on the utility scale and values on the scale of subjective probability are unrelated. It should be noted that decision-theoretical models have, to date, been mainly concerned with gambling-type situations (chance-related) which, in the language of decision theory, are situations of risk rather

than uncertainty (the assumption being that personal endeavor cannot affect the outcome). Decision situations where personal endeavor can affect the outcome (as in our ego-related situation) are, however, important and common situations in everyday life and need detailed analysis. Recently Drèze (1958) has attempted to extend decision theory in this direction. It is worth noting that the results of the present study suggest that the independence assumption may be an oversimplification for chance-related situations, achievement values related to difficulty also generalizing to these. Atkinson (1957, pp. 370–371), in his analysis of risk-taking behavior, makes a similar point.

The inverse relationship also helps to clarify an early hypothesis investigated by Wright (1937) and followed up in a number of related studies (Irwin, Armitt, & Simon, 1943; Child, 1946; Filer, 1952) that the interpolation of a barrier tends to enhance a positive valence. It places Wright's proposition in an operational framework by specifying some of the conditions under which "distant fields" may be assumed to look "greener." These conditions arise when "distant" is defined in terms of success probability, when the "greenness" refers not to the fields per se but to their attainment, and as the situation becomes more ego-related and achievement-oriented.

The direct relationship of Hypothesis 2 is in line with the assumed constraining effect of subjective probability in the Lewin et al. (1944) and Atkinson (1957) models, with the effect of expectation in Tolman's (1955, p. 320) equation, with Rotter's (1954, p. 108) direct relationship between behavior potential and expectancy, and with the assumed effect of subjective probability in the decision models of Coombs and Beardslee (1954, p. 262) and Edwards (1955, p. 201). In the present analysis the relationship is seen as complex, since choice potential is also assumed to be related to achievement values which are taken as a function of the success probability associated with the goal object, and situational context. Hence, as the results show, the relationship between choice potential and success probability is affected by the type of situation in which the choice is made, and, in fact, under achievement-oriented conditions, committed choice is seen to move in the direction of wishful choice.

In conclusion, the results of this study show the complex role of success probability in the defined choice situation, and underline the dangers of formulating theories of choice behavior *in vacuo* without regard to the type of situation in which the choice occurs, or the way in which the choice is structured, particularly the extent to which it involves commitment.

Summary

The effect of varying *S*'s subjective probability of attainment in a choice situation involving different goal objects was determined for ego-related vs. chance-related situational conditions, and for achievement-oriented vs. relaxed situational conditions, with attainment attractiveness (defined in terms of *S*'s stated wishes with respect to attainment, in a situation free from commitment) and choice potential (defined by a stated choice in a situation implying commitment) as dependent variables.

A factorial design was employed in which the goal objects were presented by the method of paired comparisons, with the following major results: (*a*) The attainment attractiveness of a goal object is inversely related to the associated success probability for all conditions combined; this relationship is more apparent in ego-related and achievement-oriented situations than in chance-related and relaxed situations. (*b*) The choice potential associated with a goal object is directly related to the associated success probability for all conditions combined. This relationship is in the opposite direction for the achievement-oriented situation; no significant differences are found between ego-related and chance-related situations for this relationship. (*c*) The reasons given by *S*s for their selections are generally consistent with the assumptions involved in the hypotheses. (*d*) An analysis of choice in an actual situation, when compared with the previous selection in the main experiment, shows a tendency to change in the direction of choice of the more likely candy, but where choice potential is the common dependent variable

this depends upon whether the situation is achievement-oriented or relaxed.

The results are related to current theoretical analyses involving similar concepts. It is suggested that theories of choice behavior should take account of the situational context in which the choice occurs and the extent to which the choice involves commitment.

REFERENCES

ATKINSON, J. W. Exploration using imaginative thought to assess the strength of human motives. *Nebraska symposium on motivation*. Univer. Nebraska Press, 1954.

ATKINSON, J. W. Motivational determinants of risk-taking behavior. *Psychol. Rev.*, 1957, **64**, 359–372.

CHILD, I. L. Children's preference for goals easy or difficult to obtain. *Psychol. Monogr.*, 1946, **60** (4), (Whole No. 280).

COOMBS, C. H., & BEARDSLEE, D. On decision-making under uncertainty. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.) *Decision processes*. New York: Wiley, 1954. Pp. 255–285.

DRÉZE, J. H. Individual decision making under partially controllable uncertainty. Unpublished doctoral dissertation, 1958.

EDWARDS, W. The theory of decision making. *Psychol. Bull.*, 1954, **51**, 380–417.

EDWARDS, W. The prediction of decisions among bets. *J. exp. Psychol.*, 1955, **50**, 201–214.

FEATHER, N. T. Subjective probability and decision under uncertainty. *Psychol. Rev.*, 1959, **66**, 150–164.

FILER, R. J. Frustration, satisfaction, and other factors affecting the attractiveness of goal objects. *J. abnorm. soc. Psychol.*, 1952, **47**, 203–212.

GUILFORD, J. P. *Psychometric methods*. (2nd ed.) New York: McGraw-Hill, 1954.

IRWIN, F. W., ARMITT, F. M., & SIMON, C. W. Studies in object-preferences I. The effect of temporal proximity. *J. exp. Psychol.*, 1943, **33**, 64–72.

LEWIN, K. The conceptual representation and measurement of psychological forces. *Contr. psychol. Theory*, 1938, **1**, No. 4.

LEWIN, K. Defining the 'field at a given time'. *Psychol. Rev.*, 1943, **50**, 292–310.

LEWIN, K., DEMBO, T., FESTINGER, L., & SEARS, P. S. Level of aspiration. In J. McV. Hunt (Ed.) *Personality and the behavior disorders*. New York: Ronald, 1944. Vol. 1, Ch. 10.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.

PHARES, E. J. Expectancy changes in skill and chance situations. *J. abnorm. soc. Psychol.*, 1957, **54**, 339–342.

ROSS, R. T. Optimum orders for the presentation of pairs in the method of paired comparisons. *J. educ. Psychol.*, 1934, **25**, 375–382.

ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.

SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.

SUTCLIFFE, J. P. A general method of analysis of frequency data for multiple classification design. *Psychol. Bull.*, 1957, **54**, 134–137.

TOLMAN, E. C. Principles of performance. *Psychol. Rev.*, 1955, **62**, 315–326.

WRIGHT, H. F. The influence of barriers upon strength of motivation. *Contr. psychol. Theory*, 1937, **1**, No. 3.

(Received October 3, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

STIMULUS CORRELATES FOR THE JUDGED ILLUMINATION OF A SURFACE [1]

JACOB BECK

University of Pennsylvania

A widely accepted view makes the evaluation of illumination (conscious or unconscious) a necessary condition for lightness constancy. However, explanations of constancy have been offered which do not involve, and even cast doubt on, the impression of illumination as a factor (Helson, 1943; MacLeod, 1947; Wallach, 1948). An alternative view is to suppose that the impression of illumination is not a condition for constancy, but a parallel phenomenon accompanying it. According to this view, stimuli complex enough to give rise to the perception of a surface would contain both correlates for the impression of the illumination and lightness of the surface. Correlates for surface lightness have been proposed by several investigators (Helson, 1943; Wallach, 1948). The judgment of illumination has, however, seldom been specifically studied (Galli, 1931; Gates, 1915; Kardos, 1934; Kruger, 1924). Except for the experimental work in connection with disproving Buhler's hypothesis—that the basis for the judgment of illumination is the light reflected by dust particles in the air (Katz, 1935, pp. 269–271)—little or no direct research has been done on the problem of how we judge the illumination of a surface when the light source is not visible.

The experiments to be reported explore how the judged illumination of a single surface varies with its illuminance. To avoid the complications of a complex or familiar texture, the experiments were all conducted with surfaces which may be described as either "speckled," "spotted," or "striped." These surfaces are examples of a class of surfaces which may be described as having a clearly perceptible pattern texture composed of two different lightnesses. They can be specified optically as consisting of an array of light composed of a clearly perceptible pattern of two different intensity levels. Wallach (1948) has shown that a surface color will appear even when such light arrays are restricted to two neighboring regions of different light intensities. The surface lightness in that case is determined by the ratio of the lower to the higher intensity, the higher intensity always being seen as white. The present experiments sought to determine for surfaces with a clearly perceptible pattern texture composed of two different intensities (*a*) whether *O*s can make consistent judgments of surface illumination, and (*b*) whether variables can be found in the light array which will specify *O*s' judgments of illumination.

Two hypotheses were tested: (H_1) The judged illumination of a surface is determined by the higher of the two intensities of the light coming from the surface. (H_2) The judged illumination of a surface is determined by the average intensity of the light coming from the surface.

[1] The first three experiments reported here are taken from a thesis presented to the faculty of the Graduate School of Cornell University for the degree of Doctor of Philosophy. The investigation was supported by a fellowship, MF 6201, from the Public Health Service. The author wishes to thank James J. Gibson for his encouragement and advice.

METHOD

General Procedure

The experiments required *O*s to adjust illuminance on a comparison surface until its illumination was judged as the same as that of a standard surface while simultaneously viewing each with monocular vision and a motionless head in a completely dark room. The standard and comparison surfaces differed in lightness and were placed so that neither shadows nor highlights were present. The distinctness of the pattern texture was not affected by increasing the illumination.

After being seated before the viewing apertures *O* was told: (*a*) the distinction between the lightness of a surface and its illumination; (*b*) that the illumination on the standard surface would be fixed; (*c*) that his task was to adjust the illumination on the comparison surface until it appeared equal to that of the standard; (*d*) that he should first make a gross match, and then make fine adjustments around that setting; and (*e*) that there was no time limit for the task. For each intensity level of the standard, *O* made 10 separate matches (except for Exp. IV where seven matches were made). The *E* alternately set the illumination on the comparison surface to a point obviously too low or too high. Each *O* was dark adapted for 3 min. before the experiment, and approximately 1.5 min. elapsed between each of the 10 matches made. The *O*s were given several practice judgments.

At the conclusion of the illumination matches the illumination on the standard surface remained at the value used in the experiment and the illumination on the comparison surface was set at the median of *O*'s matches. The *O* was then asked to match the lightness of the comparison and standard surfaces with a chart of 12 Hering grays which varied in reflectance from 8% to 78%. The chart was illuminated by a 25-w. bulb which was turned on after *O* viewed the stimulus surface. Four different charts were prepared, each consisting of the same 12 gray samples arranged in a different order. When *O*s made more than one lightness judgment of a surface, different charts were used. Systematic judgments of surface lightness were not obtained in Exp. IV.

Observers

In Exp. I, II, and III the *O*s were graduate and undergraduate students at Cornell University. In Exp. IV, the *O*s were undergraduate students at the University of Pennsylvania. All had normal vision with correction and were naive regarding the hypotheses. In each experiment 15 *O*s were used.

Experiments I and II

Apparatus.—The apparatus consisted of two viewing boxes fitted with iris diaphragms and placed so that *O* could look through one with each eye when placing his head in a head rest located directly in front. Transparencies placed over sheets of milk glass which were illuminated from behind served as the standard and comparison surfaces.[2] These were placed in two $3 \times 1\frac{3}{4}$-in. rectangular slots located at the back of each viewing box and were approximately 26 and 22 in. from *O*'s eyes. A light box containing a 375-w. photoflood lamp was centered directly behind the milk glass in each viewing box. The lamps were connected to two variacs which controlled the intensity of the illumination on the comparison and standard surfaces. Both variacs had graduated scales running from 0 to 100 which were covered by masks. A blue filter was mounted in front of each light box which was effective (except for variac settings which were greatly below the levels of illumination used in the experiments) in minimizing, but not in eliminating changes in the color of the light due to change of current. The viewing and light boxes were encased in wood and covered by black cloth for lightproofing. A voltage regulator was used to keep the line voltage constant.

Illumination measurement.—All measurements of illumination were made from the comparison and standard sheets of milk glass with a Luckiesh-Taylor brightness meter. The raw data of the experiments were numbers from 30 to 100 taken directly from *O*'s settings of the variac. The intensities in foot-Lamberts corresponding to these numbers were found by plotting a graph of foot-Lamberts vs. variac settings. The brightness of the milk glass in the comparison viewer was measured (at 11.5-v. intervals) between variac settings of 30 and 100. Since the percentage lumen output of a bulb as a function of voltage is independent of the age of

[2] The texture patterns of the surface were too small for the light reflected from them to be measured with a brightness meter. However, the light transmitted by the surfaces could be measured by a densitometer. For this reason, transparencies which were illuminated from behind were used in Exp. I and II.

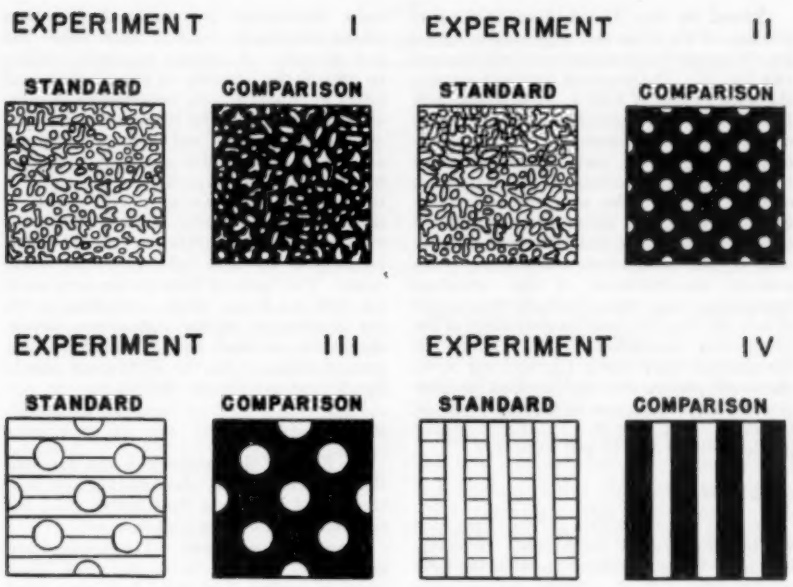


FIG. 1. Illustrations of the standard and comparison surfaces employed in Exp. I, II, III, and IV. Horizontal lines indicate gray.

the bulb, changes in the correspondence between brightness of the milk glass and variac settings resulting from aging of the bulb can be accounted for by converting the graph of foot-Lamberts vs. variac settings into a graph of percentage foot-Lamberts vs. percentage variac settings. In order to do this, the variac setting yielding 115 v. was taken as the variac setting standard, and the brightness in foot-Lamberts of the milk glass at this setting was taken as the brightness standard. Changes in the illumination of the comparison surface for different variac settings as a result of bulb aging was then determined by measuring the brightness of the milk glass at 115 v. (the standard variac setting) and multiplying it by the percentage brightness figures. The illumination of the standard surface was checked by directly measuring the brightness of the milk glass on which it was placed. After every three *O*s (except for four times after four and once after five) the foot-Lamberts corresponding to the settings of the two variacs were checked and adjustments for bulb aging made when necessary. At least five readings were taken in every measurement of the illumination.

Stimuli for Exp. I.—The transparencies used as the standard and comparison surfaces in Exp. I are illustrated in Fig. 1 and their area was approximately 50% clear and 50% dense. The standard surface appeared to *O*s as gray with white speckles, and the comparison surface appeared as black with white speckles.

The transmittances of the clear portion of the standard and comparison transparencies were the same 92%. The dense portion on the standard transparencies transmitted 35% of the light and the dense portion on the comparison transparency only 1.6%. The calculated average transmittance of the standard transparency was 1.35 times as great as that of the comparison transparency. The calculated value was checked photometrically and found to correspond to the measured average brightness. Two levels of illumination were used; the brightness of the milk glass was set at 1.1 and 3.2 ft-L. Thus according to H_1, *O* should adjust the illumination incident on the comparison transparencies to 1.1 and 3.2 ft.-L.; H_2 would predict *O* should adjust the illumination incident on the comparison transparencies to 1.49 and 4.32 ft.-L.

Stimuli for Exp. II.—In Exp. II the percentages of the dense and clear areas composing the comparison transparency were changed (see Fig. 1). The standard transparency was the same as that of Exp. I. The comparison transparency was a spotted pattern consisting of circles .04 in. in diameter and was approximately 78% dense and 22% clear. The transmittances of the dense and clear portions of both transparencies were the same as in Exp. I. Thus, the higher transmittance of the standard transparency was the same as that of the comparison transparency; the average transmittance of the standard transparency was photometrically determined to be 3.34 times the total transmittance of the comparison transparency. Two levels of illumination were used: 1.1, and 3.2 ft.-L. H_1 would predict that the illumination incident on the comparison transparency should be set by *O*s at 1.1 and 3.2 ft.-L.; H_2 would predict settings at 3.67 and 10.68 ft.-L.

Experiment III

In Exp. III, as in Exp. II, the hypothesis was tested with a spotted surface. However, instead of using transparencies with the light coming from behind, opaque surfaces were used with illumination coming from the front.

Apparatus and stimuli.—The standard and comparison surfaces were black and gray pieces of cardboard on which white circles $\frac{1}{4}$ in. in diameter had been painted. The reflectances of the gray and white, and black and white areas composing the standard and comparison surfaces were 22.5% and 83%, and 8.8% and 83%, respectively. On both the standard and comparison surfaces, the white area composed 10% of the surface. Thus, the calculated average reflectance of the standard surface was 3.26 times that of the comparison surface; the higher reflectances of both surfaces were the same.

The surfaces were placed against the back wall of a completely enclosed large rectangular box which had been divided in the center by a beaverboard partition. Two circular apertures $\frac{3}{4}$ in. in diameter were cut from the front side of the box 49 in. from the back wall. A partition placed perpendicular to the front of the box and between the two viewing apertures extended 24 in. into the room. The *O* sat with his forehead against the partition so that he could see 3×3 in. sq. areas of the comparison and standard surfaces looking with one eye through each aperture. A small fixation point was placed in the center of each of the areas.

The comparison and standard surfaces were illuminated by matched projectors placed directly in front of them below the line of sight. A voltage regulator, variacs to control the intensity of illumination, and filters to minimize color temperature changes were again used. The illumination incident on the comparison and standard surfaces was measured by taking photometer readings from a large sheet of cardboard painted with the same white as the spots and placed at their position on the back wall. Thus, all measurements of illumination referred to the intensity of the light reflected by the white spots. Two levels of illumination were used: 1.1 and 3.2 ft.-L. Thus, according to H_1 the illuminance on the comparison surface should be adjusted to 1.1 and 3.2 ft.-L., and according to H_2 the illuminance should be adjusted to 3.58 and 10.43 ft.-L.

Experiment IV

In Exp. IV the intensity of the illumination was changed by the addition and subtraction of individual light bulbs; thus, the color temperature was kept constant.

Apparatus and stimuli.—The standard and comparison surfaces were gray and black pieces of cardboard on which had been pasted three strips of white cardboard each $\frac{1}{4}$ in. wide. The area visible to *O* was 14 in. $\times$ 20 in. The white area composed 6.2% of both the standard and comparison surfaces. The gray cardboard reflected 6.4 times as much light as the black cardboard. Thus, the calculated average reflectance of the standard surface was 3.61 times that of the comparison surface. The higher reflectances of both surfaces were the same.

Illumination on the standard and comparison surfaces was provided by two light boxes, each consisting of 100 6-w. bulbs. By operating a motor which moved a beveled plate against a series of microswitches, *O* was able to turn on and off the individual bulbs of the light box illuminating the comparison surface. The box was wired so that the intensity could be varied in 40 steps from .33 ft.-c. to 12.35 ft.-c. The sound of the motor and of the opening and closing of the microswitches was muffled by padding.

The standard and comparison surfaces were set 93 in. from the light boxes in two adjacent lightproof compartments which had been divided in the center by a heavy cloth partition. A partition placed perpendicular to the front of the compartments extended 20 in. into the room. In addition to seeing the surfaces, *O*s could also see flat black curtains hanging on a wall 62 in. away.

TABLE 1

MEDIANS OF *Os* ILLUMINATION MATCHES

Measure	Exp. I		Exp. II		Exp. III		Exp. IV
	1.1 ft.-L.	3.2 ft.-L.	1.1 ft.-L.	3.2 ft.-L.	1.1 ft.-L.	3.2 ft.-L.	2.88 ft.-c.
Median illum. matches	1.15	3.21	1.17	3.29	1.30	3.06	1.89
Confidence intervals	.959–1.35	2.73–3.56	1.10–1.42	2.64–4.06	.990–1.73	2.59–3.46	1.59–2.81
Range	.840–1.45	1.60–3.65	.829–1.72	1.84–4.35	.710–2.01	1.84–4.87	1.32–3.38

All measurements of illumination again refer to the intensity of the light reflected by the white stripes. The standard surface illumination was 2.88 ft.-c. Thus, according to H_1, the illuminance on the comparison surface should be adjusted to 2.88 ft.-c.; and according to H_2 the illuminance should be adjusted to 10.4 ft.-c. All measurements of illumination were taken with a Macbeth illuminometer.

RESULTS

Table 1 shows for each experiment and condition the median for the sample, a 96.5% confidence interval for the sample, and the range of medians for individual *Os*. The confidence interval was obtained by taking the fourth and twelfth of individual *Os*' medians arranged in order of magnitude (Dixon & Massey, 1957, p. 462).

Inspection of the table indicates that: (*a*) *Os* are able to make consistent judgments of the illumination of single surfaces with clearly perceptible pattern textures, and (*b*) changing the illuminance affects *Os*' judgments of illumination. The individual data showed that all *Os* adjusted the illuminance on the comparison surface higher when the standard surface was illuminated with 3.2 ft.-L. than when illuminated with 1.1 ft.-L. Table 2, however, shows that changing the illuminance had little if any affect on *Os*' judgments of lightness. Except for Exp. III, the medians of the reflectances of the grays on the Hering chart which *Os* matched as equal in lightness to the gray and black areas of the standard and comparison surfaces remained the same for the two levels of illumination used. A median test of the differences in lightness obtained in Exp. III showed that the P values for the standard and comparison surfaces were $P < .05$ and $P > .10$, respectively.

Sign tests comparing *Os*' judgments in both Exp. I and II, and II and III

TABLE 2

MEDIANS OF THE REFLECTANCES MATCHED BY *Os* AS EQUAL IN LIGHTNESS TO THE STANDARD AND COMPARISON SURFACES

Surface	Exp. I		Exp. II		Exp. III	
	1.1 ft.-L.	3.2 ft.-L.	1.1 ft.-L.	3.2 ft.-L.	1.1 ft.-L.	3.2 ft.-L.
Standard	25%	25%	32%	32%	32%	37%
Comparison	8%	8%	8%	8%	19%	25%

Note.—Systematic judgments of lightness were not obtained in Exp. IV since only one level of illumination was used.

TABLE 3

DISTRIBUTION OF *Os* AMONG ALTERNATIVE HYPOTHESES

	Exp. I	Exp. II	Exp. III	Exp. IV
H_1	7	4	3	8
H_2	1	0	0	0
Neither H_1 nor H_2	7	11	12	7

were not significant. One can therefore conclude that the judged illumination of a smooth speckled surface is not affected (within the limits tested) by the percentage of light and dark areas or by whether the light is physically reflected or transmitted.

Test of hypotheses.—The assumption is made that if either H_1 or H_2 is true, then the values predicted should be the medians of the judgments made by each *O*. To test whether an *O*'s medians differ significantly from the predicted values a confidence region was constructed. The probability is 97.9% that an *O*'s true median will be contained within the interval formed by taking the second and ninth of an *O*'s 10 judgments arranged according to magnitude (Dixon & Massey, 1957, p. 462). Given two independent sets of 10 judgments each, the probability is therefore 98.4%, or $(.979)^2$, that the true medians of these sets will be contained within their respective confidence intervals. In Exp. I, II, and III, H_1 or H_2 was accepted if the values which they predicted were contained within an *O*'s confidence intervals. (To take an example: In Exp. I, H_1 was accepted for O_1 and H_2 was rejected because both predicted values of H_1 (1.1 and 3.2) were contained within the confidence intervals obtained by taking the second and ninth of *O*'s ordered judgments (.979–1.40, and 2.54–3.85) while the values predicted by H_2 (1.49 and 4.32) were not. Both H_1 and H_2 were rejected for O_2 because neither 1.1 and 3.2 nor 1.49 and 4.32 were contained within their respective confidence intervals (.750–.979, and 1.40–2.15).) The test is very favorable to H_1 and H_2 since it minimizes the Type I error. A similar test with a probability of 98.4% was constructed for Exp. IV by taking the first and seventh of *O*'s ordered judgments.

Table 3 shows that H_1 was accepted 7 times in Exp. I, 4 times in Exp. II, 3 times in Exp. III, and 8 times in Exp. IV; H_2 was accepted once in Exp. I and in none of the other experiments; both H_1 and H_2 were rejected 7 times in Exp. I, 11 times in Exp. II, 12 times in Exp. III, and 7 times in Exp. IV.

The results fail to support either H_1 or H_2. However, inspection of Table 1 indicates that *Os*' judgments tend to fall closer to the values predicted by H_1 than H_2. The sample confidence intervals in Exp. I, II, and III contain the values predicted by H_1 but not by H_2.

In Exp. I, III, and in particular, IV, the median illumination matches made by several *Os* were significantly below the objective illumination of the standard surface. A possible explanation is that the enhanced impression of illumination occurred because of the greater lightness differences present on the comparison surfaces. Support for the possibility of the operation of a kind of "illumination contrast" was given by two *Os* in Exp. I. They first adjusted the illumination on the black speckled comparison surfaces significantly below that of the gray speckled standard; however, when the standard and comparison surfaces were reversed, they adjusted the illumination on the now gray speckled comparison sig-

nificantly above that of the black speckled standard.

Discussion

Possible correlates for the judged illumination of the surfaces investigated are: (*a*) the higher intensity coming from the surface, (*b*) the lower intensity, (c) the average intensity, (d) the total intensity, (e) a weighted average of the higher and lower intensities. All change with surface illuminance and can serve as indicators of it. (Since the areas of the standard and comparison surfaces were the same, matches based on equating the average intensities would be identical to matches based on equating the total intensities and the experiments cannot distinguish between them.)

The results are not consistent with any of the first four hypotheses. (A match based on equating the lower intensities would have required *O* to raise the illumination on the comparison surfaces considerably above the values predicted by H_2 while *O*s' actual matches consistently fell below the values predicted by H_2.) Generally, *O*s' judgments tended to fall closer to the values predicted by H_1 than H_2. Individual differences which occur make it difficult to determine the basis for *O*'s judgments. They may have been determined by a weighted average of the lower and higher intensities. However, personal biases (a tendency to over- or undermatch), and the fact that adaptation could be only roughly equated may account for some of the deviations from H_1. No consistent trends in *O*s' judgments were observed with the time spent in the dark. Another possibility is that there is no single correlate for all *O*s. Since a number of stimuli may indicate illuminance, different *O*s may respond differently to them.

The results of the experiments need to be qualified. It is likely that with different kinds of surface textures, different variables will correlate with judged illumination. Another limitation is that the surfaces used possessed no microstructure and, in addition, had been placed so that neither shadows nor highlights were present. Changes in illumination were thus indicated only as a change in the *brightness* of the surface, i.e., a change in its intensive appearance. Ordinarily, however, illumination is also indicated by highlights, shadows, edge gradients, etc. (Evans, 1948). The matter is further complicated by the fact that judgments of surface illumination also seem to be influenced by the spatial arrangement of neighboring surfaces in the visual field (Kardos, 1934). Consequently, the results reported cannot be generalized beyond the simplified conditions of the experiments.

The results suggest the possibility of independently specifying the correlates for both the judged lightness and illumination of a surface. Kardos (1929) makes the assumption that only when visual experience is doubly determined (correlates for both judged illumination and lightness are present) can *O* distinguish between a change in the illumination and a change in the reflectance of a surface. The view suggested is that the judgments of lightness and surface illumination are direct corollaries of the perception of a surface. Stimuli complex enough to give rise to the perception of a surface will contain correlates for the judgment of both the lightness and illumination of the surface. A knowledge of these correlates would permit specification of those conditions where changes in either the objective illumination or the reflectances of the surfaces in the visual field would (*a*) leave judgments of lightness constant (lightness constancy), (*b*) leave judgments of illumination constant (illumination constancy), or (*c*) change both judgments of lightness and illumination.

Summary

A series of experiments studied an *O*'s ability to compare the illumination of two surfaces of different lightness. The *O* was given the task of adjusting the illuminance on a comparison surface until its illumination was the same as that of a standard surface. A binocular matching technique using an adjustment method was employed with the

standard and comparison surfaces being viewed simultaneously, each with monocular vision and a motionless head, in a completely dark room. Two hypotheses were tested: (*a*) The perceived illumination of a surface is determined by the higher intensity of the light coming from the surface. (*b*) The perceived illumination of a surface is determined by the average intensity of the light coming from the surface. The specific hypotheses were not supported. The *Os* were, however, able to make consistent judgments of illumination.

References

Dixon, W. J., & Massey, F. J. *Introduction to statistical analysis.* New York: McGraw-Hill, 1957.

Evans, R. M. *An introduction to color.* New York: Wiley, 1948.

Galli, A. Contributo allo studio della percezione delle variazioni di chiarore e di illuminazione. *Contr. Del Lab. Di Psychol.*, 1931, **5**, 123–143.

Gates, E. J. On intensive and qualitative judgments of light sensations. *Amer. J. Psychol.*, 1915, **26**, 296–299

Helson, H. Some factors and implications of color constancy. *J. opt. Soc. Amer.*, 1943, **33**, 555–567.

Kardos, L. Dingfarbenwahrnehmung und Duplizitatstheorie, *Z. Psychol.*, 1929, **108**, 240–314.

Kardos, L. Ding und Schatten, *Z. Psychol., Ergbd*, 1934, No. 23.

Katz, D. *The world of color.* London: Kegan Paul, 1935.

Kruger, K. Über die Unterscheidsempfindlichkeit für Beleuchtungseindrücke. *Z. Psychol.*, 1924, **96**, 58–67.

MacLeod, R. B. The effects of artificial penumbrae on the brightness of included areas. *Miscellanea Psychologica, Albert Michotte.* Louvain Inst. Sup. Philos., 1947.

Wallach, H. Brightness constancy and the nature of achromatic colors. *J. exp. Psychol.*, 1948, **38**, 310–324.

(Received September 26, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

TIME ESTIMATES AS AFFECTED BY MOTIVATIONAL LEVEL, GOAL DISTANCE, AND RATE OF PROGRESS[1]

ROBERT D. MEADE

Trinity College

Most studies of time estimates have ignored the possible effects of motivational conditions on this process. William James (1890) suggested that motivation to have time pass influences the speed of apparent time passage, but he reports no experimental data on this question. Filer and Meals (1949) have offered support for James' contention, but very few other studies have attempted to examine the role of motivation on subjective time.

A study by Hindle (1951) attempted to demonstrate that time estimates are directly related to perceived distance travelled through a task and inversely related to perceived rate of progress. In testing her hypotheses, Hindle assumed that as an *S* proceeds through a task, he perceives his rate of progress to be increasing. Thus, with this assumption, time estimates do not increase as rapidly near the goal as they do far from the goal. One aspect of the present investigation was the testing of this hypothesis in what was believed to be a more direct way.

The present investigation also attempted to handle two additional problems which were suggested by the Hindle study. Concerning the first of these, Hindle spoke of her *S*s as proceeding through a task toward a goal and assumed that her equation applied only where her *S*s had some definite end point to the task on which they were working. The use of the term "goal" would seem to imply that motivational factors were involved. The following formulation will attempt to explain how motivation to reach the end points of a task could be involved in Hindle's study.

In a situation where *S* is motivated to reach the end of a task, each unit completed will be regarded as being a significant advance toward that goal and he will see himself as making progress. On the other hand, in a situation where *S* is not motivated to reach the end of the task, each completed unit will not be regarded as having significance for him. Consequently, *S*s under the latter condition will have no appreciation of progress toward the end of the task. It is therefore predicted that time estimates will be unaffected by changes in rate of locomotion toward the end of a task if there is no motivation to reach the end of that task.

Secondly, Hindle predicted that the effect of rate of progress would be noticeable only for the latter portion of the task where *S*s have moved close enough to the goal for it to become an effective factor. This prediction, she said, depends on the assumption that a goal is more attractive the closer *S* is to it. Experiments such as those of Ovsiankina (1928), Brown (1942a, 1942b), and

[1] Based on a dissertation in psychology presented to the Graduate School of the University of Pennsylvania in partial fulfillment of the requirements for the degree Doctor of Philosophy. The writer is indebted to Francis W. Irwin for his valuable assistance during the conduct of this investigation.

Fajans (1933) have shown that certain properties of behavior vary as a function of distance from a goal. The presence or absence of motivation to reach a goal will determine whether there is a gradient of attractiveness for that point in the task. Consequently, it was predicted that time estimates will be unrelated to goal distance when there is no motivation to reach the end of a task and directly related to goal distance when there is goal motivation.

In summary, the following hypotheses were tested. When objective time is constant: (*a*) perceived duration for *S*s not motivated to reach the end of the task is unrelated either to perceived rate of locomotion or to distance from the end of the task; (*b*) perceived duration for *S*s motivated to reach the end of the task is inversely related to perceived rate of locomotion and directly related to distance from the goal.

Method

Design

This experiment used a factorial design with three independent variables—perceived rate of progress, strength of motivation to finish a task, and goal distance. There were two levels of the progress and motivation variables and three levels of goal distance. The dependent variable was estimate of time on a task when the objective time was always 6 min.

Subjects

The *S*s were 120 male college students from Trinity College and the University of Pennsylvania.

Apparatus

The apparatus was a stylus maze with 45 blind alleys and 25 choice points, a blindfold, and an electric timer. In addition, for reasons given below, several other pieces of psychological apparatus were grouped around the laboratory room, on a table and on the window sill.

Control of Independent Variables

Motivation.—An attempt was made to bring about two levels of motivation through administration of different sets of instructions to two experimental groups. In the low-motivation group, *S*s were told that they would have to spend over two hours working on tedious, routine, and boring tasks, and that nothing they did would shorten the total time they would have to spend. They were further told that no matter how well they performed, the total time could not be shortened. It was assumed that instructions of this kind would produce a low degree of motivation, or lack of motivation, to finish the first task, which was the only one they actually worked on. Subjects in the high-motivation group were told the same thing as *S*s in the low-motivation group, except that they were told that if they performed well on the *first* task, they would be excused from working on most of the other tasks which were in view around the laboratory room. It was assumed that motivation to reach the end of the first task would be high after these instructions. At the conclusion of the experiment, each *S* checked a rating scale intended to give *post hoc* indication of his motivation to finish the first task. Only those *S*s who checked that they "did not care" whether they finished the task were retained in the low-motivation group and only those *S*s who checked that they "wanted to finish very badly" were retained in the high-motivation group. These discarded *S*s were then replaced with others who met the stated criteria.

Rate of progress.—All *S*s were told that there was a given number of regions in the maze on which they were to work while blindfolded and that each of these regions could be entered in several ways. Rate of progress was controlled by giving *S* information verbally about how he was advancing from region to region in the maze. After each successive trial, *S*s in the rapid progress group were told that they had advanced either one or two regions beyond their performance on the previous trial. These scores, as well as the scores in the low-motivation group were contrived and prearranged by *E* and did not represent *S*'s true performance. They were used only to suggest to *S* how he was progressing to the goal. The *S*s in the slow progress group were told on half of their trials that they had advanced one region further and on the other half that they had advanced no further. For the last three trials of this group, *S*s were told that they

had gone no further than on the previous trial. In addition, on alternate trials, *Ss* were told that they were progressing either rapidly or slowly, depending on the group to which they belonged. In this way, it was intended that *Ss* would perceive themselves moving either at a slow or rapid rate toward the end of the task.

Goal distance.—It was intended that at the end of the experimental session, *Ss* would have traversed either 20%, 50% or 80% of the experimental task, thus leaving 80%, 50% or 20% of the total distance remaining.

In order to cause *Ss* to appear to go the same distance in the same time but at different rates, the total number of regions which *S* was told was in the maze and the number of regions he was told he had traversed were varied from one group of *Ss* to another. The *Ss* in some of the groups were told that the maze was differentiated into more subregions than for the other groups. Where rapid progress was to be perceived, there were more subregions in the maze but there was steady and rapid advancement through these regions; where slow progress was to be perceived, the same maze was represented as being divided into fewer subregions which were traversed with relatively slow progress. Table 1 shows the 12 conditions employed in the experiment, the total number of regions (denominators), and the number of regions traversed for each condition (numerators).

Procedure

After being given the appropriate instructions, *S* was blindfolded. He was then told that the experiment would begin as soon as the timer hand had moved around to the zero point, which would be exactly 1 min. This latter procedure was used to reduce variability of estimates by giving all *Ss* a standard time unit which they could use in later time estimates. The *S* was not told that he could use this as a standard. This procedure was adapted from the Hindle experiment.

The *S*'s preferred hand was then placed at the starting point of the stylus maze and he was given a series of 8 trials, each 30 sec. in duration. During the 15 sec. between trials, *S* was told his progress in terms of the regions that he had attained during the last trial. At the end of the 8 trials, a total of 6 min. had elapsed. The *S* was then asked, "How long does it seem to you that you have been working on this maze from the time we started until now?" This was followed by an interview in which the *S* was asked to rate his feelings with respect to his motivation to finish the first task.

TABLE 1

NUMBER OF REGIONS IN MAZE AND NUMBER TRAVERSED BY *Ss* IN EACH EXPERIMENTAL GROUP

Motivation	Progress	Goal Distance		
		Far	Intermediate	Near
Low	Slow	4/20	4/8	4/5
	Fast	8/40	8/16	8/10
High	Slow	4/20	4/8	4/5
	Fast	8/40	8/16	8/10

TABLE 2

MEANS OF TIME ESTIMATES IN MINUTES

Motivation	Progress	Goal Distance		
		Near	Intermediate	Far
Low	Slow	10.4	10.5	10.9
	Fast	10.4	10.5	10.1
High	Slow	6.8	8.1	10.4
	Fast	3.8	7.2	10.3

RESULTS AND DISCUSSION

The means of the time estimates for the 12 conditions are presented in Table 2 and the analysis of variance results are presented in Table 3.

Data relevant to the first hypothesis are found in the first two rows of Table 2. When these data are evaluated it is found that none of the values differ significantly from any of the others, as was predicted. This is taken as supporting the hypothesis that under low motivation, time estimates are unrelated to both perceived rate of progress and goal distance.

Data relevant to the second hypothesis, that under high motivation, perceived time is inversely related

TABLE 3

ANALYSIS OF VARIANCE OF TIME ESTIMATES

Variable	*df*	*MS*	*F*
Motivation (M)	1	214.17	107.08**
Goal distance (D)	2	65.75	32.87**
Rate of Locomotion (R)	1	18.80	9.40**
M × D	2	61.99	31.00**
M × R	1	8.26	4.13*
D × R	2	3.78	1.89
M × D × R	2	8.87	4.43*
Error	108	2.00	
Total	119	6.20	

* $.01 < P < .05$.
** $P < .01$.

to perceived rate of progress and diversely to goal distance, are found in the third and fourth rows of Table 2 and the analysis of variance in Table 3. There were significant first-order interactions between the motivation and rate of progress variables and between the motivation and goal distance variables. Under the high-motivation condition, mean time estimates decrease as the goal is approached and the decrease is more marked in the fast progress groups than in the low progress groups. At the same time, the first-order interaction between goal distance and rate of progress was not significant while the second-order interaction was significant, thus demonstrating the critical role of the motivation factor.

In addition, it was noted that the independent variables considered singly produced significant results. Consideration of Table 2 shows, however, that there is no significant variation between any of the values in the low-motivation condition. The significant variation occurs only between the high- and low-motivation groups and between the conditions of the rate of progress and goal distance variables in the high-motivation condition. These data indicate again that motivation to reach the goal is a necessary condition for the appearance of any effects produced by rate of progress and goal distance. These results were taken as supporting the second hypothesis.

It is proposed that where motivation to reach a goal is low or absent, cues to progress or goal proximity are irrelevant to *S*'s behavior and are ignored. However, under a condition where there is motivation to reach a goal, cues to distance and progress become important aspects of *S*'s behavior and thus affect the perceived duration, as reported by Hindle.

Woodrow (1951) has reviewed studies in time perception and has concluded that one of the outstanding characteristics of experiments on this process is the conflicting nature of the data from one study to another. However, the studies which he cites do not control motivational factors which may have confounded the results. This experiment demonstrates that motivation is a significant factor which can distort time preceptions and which may account for some of the ambiguity of data reported by Woodrow.

SUMMARY

This investigation was undertaken to test the following hypotheses: (*a*) For *Ss* not motivated to reach the end of a task, perceived duration is unrelated either to rate of locomotion or to distance from the end of the task, and (*b*) for *Ss* motivated to reach the end of the task, perceived duration is inversely related to rate of progress and distance from the goal. Half of the 120 *Ss* were given low-motivation instructions and half were given high-motivation instructions. All *Ss* spent 6 min. performing 8 trials on a stylus maze while blindfolded. After each trial, they were given contrived scores about how rapidly they were advancing toward the goal, with *Ss* finishing at three different distances to the goal. At the end of 6 min., *Ss* were asked to estimate how long it seemed that they had been working on the task. The results supported both hypotheses.

References

BROWN, J. S. Factors determining conflict reactions in difficult discriminations. *J. exp. Psychol.*, 1942, **31**, 272–292. (a)

BROWN, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226. (b)

FAJANS, S. Die Bedeutung der Entfernung dür die Stürke einer Aufforderungscharakters beim Saugling und Kleinkind. *Psychol. Forsch.*, 1933, **17**, 213–267.

FILER, R., & MEALS, D. The effect of motivating conditions on the estimation of time. *J. exp. Psychol.*, 1949, **39**, 327–331.

HINDLE, H. Time estimates as a function of distance travelled and relative clarity of a goal. *J. Pers.*, 1951, **19**, 483–501.

JAMES, W. *Principles of psychology*. New York: Holt, 1890.

OVSIANKINA, M. Die Wiederaufnahme unterbrochener Handlungen. *Psychol. Forsch.*, 1928, **11**, 302–379.

WOODROW, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1224–1226.

(Received October 1, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

INHIBITORY SET IN PROBLEM SOLVING AS RELATED TO REINFORCEMENT LEARNING

ROBERT ADAMSON[1]

Emory University

The phenomenon of *Einstellung*, or set induced by the repetitive success of a particular solution method over a series of problems, has been considered either as an individual difference variable (Cattell, Tiner, & Winder, 1949; Cattell & Winder, 1952; Cowen, Wiener, & Hess, 1953) or as a phenomenon imposed by arrangement, experimental or other. Consideration of the latter point of view suggests that methods of solution may become fixed in a manner analogous to the fixation of a CR; that is, a particular method is reinforced if it leads to solution, successive such reinforcements leading to fixation. If so, the method should persist whenever *S* encounters similar stimuli, even when the method is no longer appropriate. Such is the case in the Luchins (1942) series of water-jar problems.

This approach, S-R oriented, has been applied successfully in studies of concept formation (Buss, 1950) and of discrimination reversal (Buss, 1953). Kendler, Greenberg, and Richman (1952), with a similar orientation, investigated the effect upon extinction of set of massed vs. distributed presentation of set-inducing problems, finding that massed presentation produced greater resistance to a shift to a new method of solution.

The present study views set-induction in the context of general reinforcement learning; consequently, an important variable in determining the persistence of a particular method should be its reinforcement history. In accord with this approach, the hypothesis to be tested is that intermittent success of a specific method of anagram solution—analogous to partial reinforcement—will induce greater resistance to extinction of that method than will every-trial success.

METHOD

Subjects.—A total of 32 *Ss*, undergraduate students at Emory University, were placed into two groups having respective *Ns* of 12 and 20. These groups were matched for sex.

Procedure.—The task, presented individually to *Ss*, was that of solving a series of four-letter anagrams, successively projected on a screen in front of *S*. Each *S* was shown an example anagram (GRWO) which he solved before proceeding with the test series. He was instructed to close a switch as soon as he had thought of a solution. This switch turned off the slide projector and also broke a timing circuit. As soon as *S* had closed the switch, he wrote down his solution, following which the next slide was presented. Thus, there was a record of the correctness of each solution and one of time-to-solution for each anagram.

The anagram series consisted of three groups: set-induction, extinction, and reinstatement. For one group of 12 *Ss*, the partial reinforcement (50%) group, the set-induction series consisted of 12 anagrams, six of which could be solved only by transposing letters in a 2–3–4–1 order (e.g., ETYP). These six were randomly interposed among six other anagrams, none of which could be solved by the foregoing, or by a common, method. For another group of 20 *Ss*, the every-trial (100%) group, the set-induction series consisted of the 2–3–4–1 anagrams presented in succession.

The extinction series followed and was common for all *Ss*. It consisted of six anagrams for which the 2–3–4–1 method was inappropriate and for which there was no method common to any two anagrams. The reinstatement series, which followed, con-

[1] Now with System Development Corporation.

sisted of four anagrams which could be solved either by the 2–3–4–1 method or by a 4–3–2–1 method (e.g., ELUR).

The anagrams used in the series were chosen from a list compiled by McNemar and Taylor (1955). The 2–3–4–1 anagrams and the other six anagrams constituting the set-induction series for the 50% group were matched for difficulty in the above study, in terms of percentage of correct responses.

Results

Erroneous solutions failed to differentiate between groups or conditions. Accordingly, the following analyses were based on the time-to-solution measure.

For the 50% group, solution scores for the set-induction series revealed that the problems having a common method were more quickly solved than were the heterogeneous problems. The mean time-to-solution for the six 2–3–4–1 problems was 3.99 sec.; for the others, it was 16.49 sec. ($t = 6.44$, $df = 10$, $P = < .01$). This could indicate a facilitative effect for the common method, or, alternatively, that solution of other anagrams in the same series is inhibited.

The method established in the set-induction series exerted an inhibitory effect on solution of the problems in the extinction series (see Table 1).

This effect was true for both the 50% group ($t = 2.94$, $df = 10$, $P = < .02$) and for the 100% group ($t = 2.86$, $df = 17$, $P = < .02$). The inhibitory effect was more pronounced for the 50% group than for the 100% group, in accord with the hypothesis. The t for the difference between mean differences (3.25) was 10.48 ($df = 27$, $P = < .001$).

Scores on the reinstatement series revealed no significant differences between the two groups, either in terms of choice of solution method or of time-to-solution. Thus, whatever differential existed during the extinction series did not carry over to the reinstatement series.

TABLE 1

Mean Time-to-Solution for Anagrams in the Set-Induction and Extinction Series

Group	N	Mean Time-to-Solution (Sec.)		
		Induction	Extinction	Diff.
50%	12	3.99	12.82	8.83
100%	20	3.94	9.34	5.58

Discussion

The results uphold the hypothesis that partial success of a particular method tends to cause greater resistance to extinction of the method than does every-trial success. By analogy, these results may be considered a partial reinforcement effect, and tend to support a research orientation allied to that of reinforcement learning in general.

More specifically, these results support the position that an important factor in determining set formation is the reinforcement history of the method itself. One might consider such a conclusion contradistinctive to that advanced by Kendler et al. (1952), for whom the important determinant of a set for a particular response was the extinction of alternative response tendencies during successive problem exposures, during which there was no selective strengthening of the specific method. Their experiment, however, used as a measure of extinction performance on a single problem which could be solved in either of two ways: by the previously successful "long" method or by a direct method. Designating such a performance a measure of extinction, as had also been done by Luchins (1942), is to translate a term from its conventional meaning to a new one which is analogously inappropriate, since in extinction the making of a previously learned response is no longer reinforced (or successful). Such is the case in the extinc-

tion series of the present experiment, but not in the above-cited studies (Kendler et al., 1952; Luchins, 1942).

If sets are formed according to conventional reinforcement principles, further experimentation relating to the amount of work involved, the effects of rest intervals following extinction, etc., should yield results consonant with the preceding in terms of orientation.

Summary

On the assumption that the development of inhibitory set proceeds according to the principles of reinforcement learning, such a process should display phenomena characteristic of reinforced responses in general. One such phenomenon is the relative success of partial reinforcement in producing resistance to extinction. Accordingly, intermittent success of a solution method should cause greater resistance to extinction of the method in inappropriate problem situations than should every-trial success.

This hypothesis was tested by presenting three series of anagrams to human *Ss*: set-induction—in which partial was opposed to every-trial success for two groups—extinction, and reinstatement. With a time-to-solution measure, the results upheld the hypothesis. It was found that intermittent success with a given solution method tended to favor its persistence in inappropriate situations relatively more than did every-trial success.

References

Buss, A. H. A study of concept formation as a function of reinforcement and stimulus generalization. *J. exp. Psychol.*, 1950, **40**, 494–503.

Buss, A. H. Rigidity as a function of reversal and nonreversal shifts in the learning of successive discriminations. *J. exp. Psychol.*, 1953, **45**, 75–81.

Cattell, R. B., & Winder, A. E. Structural rigidity in relation to learning theory and clinical psychology. *Psychol. Rev.*, 1952, **59**, 23–39.

Cattell, R. B., Tiner, L. G., & Winder, A. E. The varieties of structural rigidity. *J. Pers.*, 1949, **17**, 321–341.

Cowen, E. L., Wiener, M., & Hess, J. Generalization of problem-solving rigidity. *J. consult. Psychol.*, 1953, **17**, 100–103.

Kendler, H. H., Greenberg, A., & Richman, H. The influence of massed and distributed practice on the development of mental set. *J. exp. Psychol.*, 1952, **43**, 21–25.

Luchins, A. S. Mechanization in problem solving: The effect of *Einstellung*. *Psychol. Monogr.*, 1942, **54** (6), (Whole No. 248).

McNemar, O. W., & Taylor, D. W. Anagram and puzzle series for the study of set. Stanford: Stanford Univer., 1955 (Cont. N6onr 25125).

(Received October 10, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

AMOUNT OF CONDITIONING AND INTENSITY OF CONDITIONED STIMULUS[1]

HERBERT D. KIMMEL[2]

University of Southern California

This study evaluated the hypothesis that amount of classical conditioning, measured during the acquisition process, would vary as a function of the intensity of the conditioned stimulus (CS). The direction of the hypothesized relationship was not predicted but it was expected to be nonlinear and possibly not monotonic.

The experiment was done in the context of Razran's (1957) theory of the acquisition of classical conditioning. Razran's theory predicts that a positive relationship between CS intensity and amount of conditioning would hold in the lower portion of the CS intensity continuum but that the relationship would become negative when the value of the CS-initiated neural event approached that of the unconditioned stimulus (UCS)-initiated neural event and the latter no longer "dominated" the former (1957, p. 13).

Previous American studies of the role of CS intensity in classical conditioning (Carter, 1941; Grant & Schneider, 1948, 1949; Hovland, 1937) have reported negative results. All but Carter (1941), however, evaluated differences in amount of conditioning during extinction trials. For this reason, their findings cannot be considered applicable to Razran's theory which is concerned exclusively with the acquisition process.

The present study was designed to permit amount of GSR (to a 1000 cps tone) on acquisition trials to be used as a measure of conditioning. To provide support for the assumption that observed differences could be attributed to conditioning effects, some *Ss* were used as controls who did not receive paired presentations of the CS and UCS but were treated similarly to the experimental *Ss* in other respects.

METHOD

Subjects[3] *and design.*—The *Ss* were 72 volunteer undergraduate students at the University of Southern California. They were given a hearing test to ensure that none had more than a 15-db hearing loss in either ear. The *Ss* were assigned randomly to two equal-sized groups. The groups differed with respect to the presence or absence of a paired relationship between CS and UCS; the group that received paired presentations is referred to below as the *conditioning group* (Group C); the group that received unpaired presentations of CS and UCS is referred to as the *pseudo-conditioning group* (Group P). Each group was divided randomly into three subgroups of 12 *Ss* each. The three subgroups differed with respect to the CS intensity used; one conditioning subgroup and one pseudo-conditioning subgroup received a CS intensity of 35 db, 75 db, and 115 db, respectively (all db re .0002 dynes/cm^2).

Instructions.—All instructions were administered in writing except as noted below. Three separate written instructions were given. The first explained the nature of the GSR and the stimuli to be used. It indicated that this was an experiment on the way in which different stimuli evoke the GSR. The second written instruction explained how the shock intensity was to be set. The third

[1] This paper is based on a portion of the data of a Ph.D. dissertation submitted to the Graduate School of the University of Southern California. W. W. Grings was chairman of the dissertation committee.

[2] Now at the University of Florida.

[3] The *Ss* were obtained from a subject-pool made possible by a grant from the National Science Foundation to W. W. Grings, Psychology Department, University of Southern California.

written instruction told *S* to relax his body and remain motionless and it explained that a blindfold was to be used to eliminate visual stimulation.

Apparatus.—The CS was a 1000 cps tone generated by a Hewlett-Packard Model 200-B audio oscillator. The oscillator was on at all times during the data collection period. The intensity of the tone was determined by the setting of a Hewlett-Packard Model 350-B attenuator as follows: With the attenuator set at zero attenuation, the output of the oscillator was manipulated until the voltage across the PDR-8 headphones (with MX-41/AR ear cushions) was 1 v. The headphones had been calibrated beforehand to determine that each phone delivered approximately 115 db when 1 v. was measured across it. The oscillator setting that produced the desired voltage was maintained throughout the experiment. Variations in CS intensity by *E* during the experiment were accomplished by changing the attenuator setting.

The UCS was an electric shock delivered to the volar surface of *S*'s forearm (preferred) through two ¾-in. silver electrodes spaced 1 ¼ in. apart. The shock was generated by the single discharge of a .1-mf. condensor, charged to 90 v. and discharged through the primary of a 1:18 transformer. Intensity was varied by a voltage divider in the secondary of the transformer. The strength of the shock was adjusted for each *S* to meet a double criterion: the consistent production of a GSR and an oral report of definite annoyance. This setting was achieved by presenting *S* with a barely perceptible shock and increasing the intensity in small steps until *S* reported definite annoyance.

Temporal relations between the CS and UCS were controlled by two Hunter electronic timers. Group C received a tone of 4 sec. duration followed immediately by the electric shock on acquisition trials. Group P received the 4-sec. tone alone and the shock alone as described below. The timers also drove two signal-magnet pens indicating onset and cessation of the tone and shock on the GSR record.

The GSR was picked up by a saline paste-metal-skin junction using ¾-in. Lucite cups containing zinc electrodes and fastened by rubber straps to the palm and back of *S*'s nonpreferred hand. Alcohol was used to clean *S*'s hand prior to attaching the electrodes. The response was measured as a resistance change in a Wheatstone bridge input to a Brush DC amplifier, with changes in resistance recorded as deflections on an Esterline-Angus ink-writing milliammeter with a paper speed of 3 in. per min. The current through *S* was maintained at 42.5 microamperes.

Procedure.—Observations were made in a soundproof room with all lights out except for a covered 7.5-w. lamp in *E*'s cubicle. A partition separated *S* from *E* and the equipment. The oscillator, timers, and associated relays were in another room. The only noise in *S*'s room was the click of the signal-magnets on the recorder and the hum of an electric fan masking these clicks.

At the beginning of the data collection period, *S* was seated and given a hearing test with a Beltone audiometer. Then *S* read the first written instruction. After *E* had attached the GSR and shock electrodes to *S*, the second written instruction was read by *S*. Meanwhile *E* checked the calibration of the GSR record. Then the intensity of the shock was set and *S* was told that it would remain the same for the remainder of the experiment. The *S* then read the third written instruction and was told to relax and pay attention to the stimuli. The headphones and blindfold were attached and the lights turned out.

All *S*s received six preliminary trials of tone alone. The CS intensity was used for three of these preliminary trials and a different intensity (not one of the three CS intensities) was used for the remaining three trials. A counterbalanced order was used for the two intensities presented on preliminary trials. The preliminary trials were followed by three presentations of the shock alone. These presentations were followed by six adaptation trials with the CS alone. The preliminary trials, shock alone trials, and adaptation trials were all separated by 20–30 sec. unsystematically varied by *E*.

The *S*s in Group C then received 20 acquisition trials of tone followed by shock. These trials were separated by 40–60 sec. unsystematically varied by *E*. The *S*s in Group P received 20 presentations of the CS and 20 shocks in an order designed to prevent the formation of patterning sets. These trials were separated by 20–30 sec. unsystematically varied by *E*.

The *S*s were then paid and asked to tell *E* what they noticed about the experiment. Before leaving they were cautioned against discussing the experiment with others.

Results

Amount of GSR was measured as the maximum deflection of the re-

corder pen within a 5-sec. interval beginning with the onset of the CS. The unit employed was the square root of conductance change ($\sqrt{\Delta C}$), which has been found in previous research (Grings & O'Donnell, 1956) to be distributed appropriately for the statistical procedures used. The 5-sec. interval excluded unconditioned responses to the UCS (GSR latency is almost never less than 1 sec.).

The acquisition score was obtained by subtracting the average $\sqrt{\Delta C}$ to the CS on the three preliminary trials from the $\sqrt{\Delta C}$ measured on each acquisition trial. Acquisition scores from pairs of adjacent trials (i.e., 1st and 2nd, 3rd and 4th, etc.) were summed to increase their stability. There were thus 10 acquisition measures for each *S*. The subtraction process was used to control for differences in unconditioned response strength of the different CS intensities. It also had the effect of reducing differences among *S*s in GSR responsiveness.

The distributions of acquisition scores for Groups C and P were quite symmetrical. The reliability of these scores was estimated approximately by obtaining the correlation of the sum of responses on odd-numbered trials and the sum of responses on even-numbered trials. Corrected for double length this correlation was found to be .81 (all *S*s). On the basis of the observed symmetry and reliability of the acquisition scores it was concluded that they were appropriate for the analyses presented below.

To evaluate the experimental hypothesis that differences in amount of conditioning would be a function of CS intensity, a separate comparison of trial-by-trial acquisition scores was made for each pair of conditioning and pseudo-conditioning subgroups receiving the same CS intensity. These comparisons are presented graphically in Fig. 1.

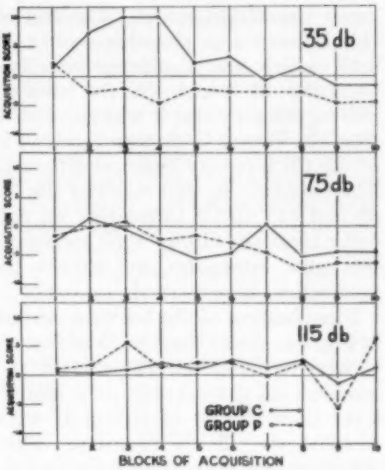


FIG. 1. Acquisition score for the GSR in successive blocks of two trials for the conditioning (Group C) and pseudo-conditioning (Group P) groups, with three intensities of the CS.

The figure indicates that Group C *S*s who received a CS intensity of 35 db behaved quite differently from Group P *S*s receiving the same intensity. The Group C subgroup showed a gradual increase in average acquisition scores up to the fourth block of two acquisition trials, reflecting a conditioning effect, while the Group P subgroup showed a steady decrease in average acquisition scores, reflecting an adaptation effect. No differences appeared to occur between Group C and P *S*s receiving the 75 db and 115 db CS intensities, indicating that conditioning did not occur for these intensities.

The significance of the observed difference between Group C and P *S*s receiving a CS intensity of 35 db was evaluated by comparing their over-all

mean acquisition scores. The test of this difference produced a t ratio of 4.17, which was significant well beyond the .01 level.[4] On the basis of this significance test it was concluded that the Group C *Ss* who received a 35 db CS intensity were conditioned. The Group C *Ss* who received the 75 db and 115 db CS intensities did not differ from the Group P *Ss* receiving the same intensities and cannot be assumed to have learned.

Examination of the learning curves in Fig. 1 suggests that the *Ss* of Group C who received the 115 db CS responded, on the average, at a higher level than the *Ss* of Group C who received the 75 db intensity. This difference was observed for the Group P *Ss* receiving these intensities also, indicating that it was not due to conditioning. It was possible that a sensitization effect, due to the intensity of the CS, was responsible for this difference. To check on this possibility, measures of skin conductance prior to and after the acquisition series (between trials) were obtained for each *S*. Since skin conductance has been assumed to indicate level of activation (Woodworth & Schlosberg, 1954), it was expected that the *Ss* receiving the 115 db CS would show a greater change in skin conductance over the entire acquisition series than was shown by *Ss* receiving the less intense CSs. This would support the notion that sensitization, due to the activating effect of the most intense CS, was responsible for the difference in acquisition scores between Group C and P *Ss* receiving the 115 and 75 db intensities. These conductance measures are presented in Table 1.

[4] An analysis of variance of the 720 acquisition scores (all *Ss*) provided an estimate of error variance for this t ratio (cf. Lindquist, 1956, p. 146).

TABLE 1

MEDIAN CONDUCTANCE VALUES (MICROMHOS) BEFORE AND AFTER ACQUISITION

Stimulus	Base Conductance			
	Before Acquisition		After Acquisition	
	Group C	Group P	Group C	Group P
35 db	36.6	37.1	37.8	38.4
75 db	36.8	37.7	39.0	38.5
115 db	35.6	30.2	39.0	35.3

As was expected, the amount of change in skin conductance over the entire acquisition series was greatest for *Ss* receiving the 115 db intensity and least for *Ss* receiving the 35 db intensity. Median values are used in the table because of the large ranges and lack of previous information regarding the distribution of these measures. The median values before acquisition were not significantly different from each other (tested by comparing the two most discrepant groups with the Median test), nor were they significantly different from each other after acquisition. The only subgroup showing a significant change in conductance was the Group P 115 db subgroup, which showed a significant increase (Wilcoxon Paired Replicates test, $T = -6$, $P < .01$). This significant increase, coupled with the apparent over-all trend in amount of change in skin conductance as a function of CS intensity, supported the notion that a sensitization effect, namely an increase in level of activation, was responsible for the difference between Group C and P *Ss* receiving the 115 db intensity and Group C and P *Ss* receiving the 75 db intensity.

DISCUSSION

The results of this experiment appear to support Razran's theoretical assertion

that strong CS may complicate or interfere with the formation of CR. Since the weakest intensity used in this study was associated with the most conditioning (more correctly, with the *only* significant conditioning), it would be difficult to predict from the present data whether weaker intensities would lead to more or less conditioning when compared with a 35 db CS.

A question may be raised as to why the optimum CS intensity in this study should have been so low in the auditory intensity continuum. Two speculative comments may help answer this question. The first is that the peculiar nature of the GSR, in comparison with autonomic responses such as salivation, may make it more susceptible to interference and subsequent poor conditionability. The GSR differs from salivation in being an organismic rather than a localized response, diffuse rather than specific. If it is assumed that the lack of "dominance" mentioned in Razran's theory stems from the tendency of the CS to interfere with the UCS-UCR reflex, then highly sensitive organismic responses, such as the GSR, might be conditionable more readily to stimuli that, themselves, have little unconditioned propensity to elicit the response. In this study only the 35 db CS was close to being inadequate to produce the GSR before acquisition.

The second speculation regarding the reason for the 35 db CSs superiority over the other CSs derives from a statement by Pavlov (1927) regarding CS intensity and conditioning. Pavlov observed that the defensive or investigatory reflexes evoked by any CS were definite obstacles to the formation of CRs. In the present study, it can be assumed that the 35 db CS had less of a tendency to arouse defensive or investigatory reactions than the more intense CSs and, thus, was conditioned more readily.

The finding that sensitization occurred in *Ss* receiving the 115 db CS, whether or not it was paired with the UCS, has two important implications. The first is the fact that conditioning studies employing mildly noxious CSs require adequate sensitization controls to be meaningfully interpreted. In the present study, for example, it would have been difficult to differentiate between the Group C subgroups receiving the 35 db and 115 db CSs had no control group been used. The second implication of the sensitization effect observed in this study is that it appeared to be due to the CS rather than to the UCS, as might have been expected. This probably means that sensitization controls in classical conditioning studies should receive presentations of the CS as well as the UCS to support adequately the assumption that conditioning occurred in experimental groups.

Summary

An experiment was conducted to test the hypothesis that amount of conditioning is a function of CS intensity. The experiment was done in the context of Razran's theory of classical conditioning. The palmar GSR, produced by electric shock to the forearm, was conditioned to a 1000 cps tone of three intensities: 35, 75, and 115 db (re .0002 dyne/cm^2). The CS was of 4-sec. duration and was followed immediately by the UCS on 20 acquisition trials. For each CS intensity a control group was run receiving unpaired presentations of tone and shock. Amount of conditioning was measured in terms of the amount of GSR produced by the CS on acquisition trials over and above amount of GSR produced by the CS on preliminary trials.

The results showed that conditioning occurred only in the group receiving paired presentations of the 35 db CS and the shock UCS. The more intense CSs failed to be conditioned. Sensitization occurred in the groups receiving the 115 db CS, whether or not it was paired with shock. The results were interpreted as supporting Razran's theory of classical conditioning.

References

Carter, L. F. Intensity of the conditioned stimulus and rate of conditioning. *J. exp. Psychol.*, 1941, **28**, 481–490.

Grant, D. A., & Schneider, D. E. The intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *J. exp. Psychol.*, 1948, **38**, 690–696.

GRANT, D. A., & SCHNEIDER, D. E. The intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *J. exp. Psychol.*, 1949, **39**, 35–40.

GRINGS, W. W., & O'DONNELL, D. E. Magnitude of response to compounds of discriminated stimuli. *J. exp. Psychol.*, 1956, **52**, 354–359.

HOVLAND, C. I. The generalization of conditioned responses. IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *J. exp. Psychol.*, 1937, **21**, 261–276.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1956.

PAVLOV, I. P. *Conditioned reflexes.* London: Oxford, 1927.

RAZRAN, G. The dominance-contiguity theory of the acquisition of classical conditioning. *Psychol. Bull.*, 1957, **54**, 1–46.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* New York: Holt, 1954.

(Received October 1, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

STIMULUS SIMILARITY AND TASK FAMILIARITY AS DETERMINANTS OF EXPECTANCY GENERALIZATION

DOUGLAS HEATH [1]

Haverford College

Recently, there have been attempts to understand expectancies as cognitively learned units of behavior subject to reinforcement learning principles (Chance, 1952; Jessor, 1954; Rotter, 1954). Although *S*'s expectancies may not be explicitly or precisely formulated, it has been possible to objectify the construct by defining an expectancy as that score *S* predicts he will make in a specific task, and expectancies so defined can be modified by reinforcement operations. A negatively reinforced expectancy (failure) will be lowered in subsequent trials; a positively reinforced expectancy (success) will usually be raised (Lewin, Dembo, Festinger, & Sears, 1944; Rotter, 1942).

Reinforcement of an expectancy may affect expectancies about other quite different situations. For example, a bright ninth grade student unexpectedly fails algebra. Why does he subsequently expect to fail geometry, physics, and possibly history? Rotter has called this effect an example of the generalization of expectancy changes which, he has hypothesized, is mediated more by the goal relatedness, or the similarity of reinforcement values, of the stimuli along a generalization continuum than by objective physical stimulus characteristics. Attempts to obtain a definition of "goal relatedness" for a range of stimuli independent of other stimulus properties and phenomenologically valid for each *S* have not been particularly successful (Chance, 1952; Jessor, 1954) primarily because "goal relatedness" has been defined a priori by *E* without consideration of *S*s' need hierarchies. The experimental results cited by Rotter in support of his thesis need to be replicated not only because of the experimental failure to obtain independent measures of goal relatedness, but also because of the experimental confounding of temporal order of task presentation with stimulus similarity (Jessor, 1954) and/or the use of a very restricted number of generalization tasks, e.g., one to three, at most (Jessor, 1954; Rotter, 1954, p. 121).

This paper reports two experiments that demonstrate expectancy generalization utilizing *S*-defined rather than *E*-defined criteria of stimulus similarity. To establish the fact of expectancy generalization across a range of tasks. Exp. I hypothesized that (*a*) positive and negative reinforcement of an expectancy in one task will differentially affect expectancies in related tasks, and (*b*) the degree of generalization of expectancy changes will decrease with increasing dissimilarity of the related tasks. "Increasing dissimilarity" was defined by each *S*'s postexperimental global impressionistic judgments of degree

[1] I am grateful to T. Koester, Amherst College, who suggested the study, to W. Hays and W. Norman, University of Michigan, and G. Mandler, Harvard University, for their assistance. I appreciate the support of the National Science Foundation and the Social Science Research Council which has given me time to prepare this article. The experiment was conducted at Amherst College.

of similarity among the generalization tasks. To determine if *S* performance on the generalization tests was a variable, which, by reducing the ambiguity of the situation, leads to a reduction in the amount of generalization, Exp. II hypothesized that (*c*) degree of generalization of expectancy changes will decrease with increased *S* familiarity with the task situation. "Increased familiarity" was defined by *S* performing each of the generalization tasks.

Method

Subjects.—Eighty volunteer male students in elementary psychology courses were used. The *Ss* included all but one member of the courses concerned. The *Ss* were predominantly sophomores and juniors with a minimum knowledge of psychology. The *Ss* knew *E* who assisted them in their laboratory work.

Design.—Following the instructions, each *S* stated the percentile score he expected to get on a vocabulary test which he subsequently took. The *E* then reported a falsified prearranged score that was above (success) or below (failure) *S*'s stated vocabulary expectancy. In Exp. I, *S* estimated the percentile score he expected to make on one of the generalization tests, took that test, and then gave his expected percentile score for the next generalization test. In Exp. II, *S* estimated the percentile score he expected to make on each of the six generalization tests but did not take any of the tests.

Materials.—The initial test on which *S*'s expectancy was reinforced was the Nelson-Denny Vocabulary Test, Form A, 1929 edition. The *S* selected the correct synonym from five alternative words for each of the increasingly more difficult 100 test items. Ten student judges, not part of the *S* population, were each given copies of the vocabulary test and 12 different ability tests. Each judge rated the similarity of each of the ability tests to the vocabulary test on a 3-point scale. He then ranked each test in order of its similarity to the vocabulary test. "Similarity" was deliberately left undefined. On the basis of the means and *SD*s of the judges' ratings and rankings, five tests were selected to represent a generalization continuum ranging from "very similar" to "very dissimilar" to the vocabulary test. The tests were, in decreasing order of judged similarity to the vocabulary test: (*a*) Grammar and Spelling Analogy Test, Test 2, Ohio State University Psychological Test, Form 22, 1943; (*b*) Paragraph Comprehension, from the Iowa Silent Reading Test, Form AM; (*c*) Recognition of Assumptions, from the Watson-Glaser Test of Critical Thinking, I-A, 1942; (*d*) Number Series, from the American Council of Education, 1939; and (*e*) Eye-hand Dexterity Test, from the Minnesota Spatial Relations Test. To increase the range of dissimilarity, a dart throwing test was subsequently added as a sixth test for Exp. II.

Procedure.—All *Ss* volunteered for the experiment at the same time by signing a master appointment list. The first 40 *Ss* on the list were assigned to Exp I which was completed before data were collected from the remaining 40 *Ss* for Exp. II. For each experiment, every other *S* who appeared for an appointment was routinely assigned to the failure group. Each *S* was tested individually. He was given an instruction sheet which described the experiment's ostensible purpose and procedure. The *S* was told that the experiment was measuring the accuracy of his self estimates and that "the accuracy with which an individual can judge others and himself . . . is . . . an index of intelligent behavior." He was told that he was to take a vocabulary test which was also indicative of intelligence and that he would receive a percentile score based on the college's norms. He was to state how well he would do in comparison with other men of the college. He was assured all results would be kept private and would not affect his grade in the course. The instructions and procedure were identical for the Exp. II *Ss* except that each *S* was told he would take each of the tests after he had given his estimates for all tests. After giving his estimates, the Exp. II *S* was told he did not have to take the tests. After *E* had answered *S*'s questions, *E* repeated verbally some of the written instructions. The *E* defined and gave examples of a percentile score and emphasized that *S* was to estimate the percentile score he would actually get, not what he would like or not like to get. The *S* was then given a sample of the vocabulary test after which he wrote his estimated percentile rank at the top of a sheet which had mimeographed multiple choice answers for all of the tests, except the vocabulary, dart throw, and dexterity tests, that *S* thought he was to take.

When taking the vocabulary test, *S* was told to call out the number of one of the five

answers to each word, even if he had to guess, so *E* "could correct the test as we go along." The *S* was told that the final score combined both speed and number of correct answers but that *E* could not tell him how much time he would have. At the end of 10 min., or when *S* had completed the test, *E* took the test away, performed some fake figuring, ostensibly consulted a table of norms and wrote a prearranged percentile rank at the top of *S*'s answer sheet. The same prearranged schedule of falsified scores was used for both Exp. I and II. Variable amounts were either added to (for success *S*s) or subtracted from (for failure *S*s) *S*'s initial percentile estimate, depending upon the size of *S*'s initial estimate. For example, all "failure" *S*s expecting their scores to be 20th percentile were reported by *E* to be on the 4th percentile; all "failure" *S*s expecting the 90th percentile were reported to be on the 67th percentile. Table 1, which gives the differences between the mean estimated vocabulary and reported percentile scores for each experimental condition, shows each group received about the same amount of reinforcement. Two *S*s indicated in a post-experimental interview that they doubted the validity of the reported score. These *S*s were in different experimental groups.

The *S* was next given a sample of one of the generalization tests after which he wrote his expected percentile rank on the answer sheet. He then either took the test or was told he would take it after he had given his estimates on all of the tests. Each *S* was given the tests in a prearranged but individually randomized order. No two *S*s received the tests in the same order.

At the completion of the experiment, each *S* rated on a 3-point scale the similarity of each of the tests to the vocabulary. Following the ratings, *S* then ranked each test in order of its similarity to the vocabulary. No tied ranks were allowed. Despite some requests, *E* consistently refused to give any clues about the meaning of "similarity" in order not to affect *S*'s judgments. The *S* was then interviewed about the determinants of his estimates, and was asked not to discuss the experiment with anyone else. At the end of Exp. II, *S*s were notified about the purpose of the experiment.

Results

Table 1 indicates the mean *S*-estimated vocabulary and *E*-reported vocabulary scores for each experimental condition. The differences between the vocabulary estimates of the success-failure groups for each experiment and of the success-success and failure-failure groups for both experiments were nonsignificant.

TABLE 1

Mean Estimated Vocabulary and Reported Percentile Scores

Reinforcement Condition	*N*	Vocab. % Est.	% Reported by *E*	% Diff. Score
Exp. I				
Success	20	67.75	89.0	21.25
Failure	20	61.6	39.8	21.8
Exp. II				
Success	20	61.1	86.7	25.6
Failure	20	66.1	43.5	22.6

The generalization tests did represent different degrees of dissimilarity to the vocabulary test. Mean similarity ratings for each generalization test were comparable for each experimental condition. The mean similarity ratings for the performance group were, for example: Analogy 1.05, Recognition of Assumptions 2.12, Paragraph Comprehension 2.17, Number Series 2.75, and Dexterity 2.8.

The basic data used in the analyses were *S*s' percentile estimates made for that test which was judged by each of those *S*s to have the same ranked similarity to the vocabulary test, e.g. S_1's percentile estimate for the Ohio Analogy Test which S_1 ranked 1 was combined with S_2's percentile estimate for the Number Series Test which S_2 ranked 1. By combining the percentile estimates for the individually ranked positions rather than for tests it was possible to use a *S*-defined rather than *E*- or group-defined stimulus continuum of similarity. All results, unless stated otherwise, are presented in terms of each *S*'s own judgment of the degree of similarity of the tests.

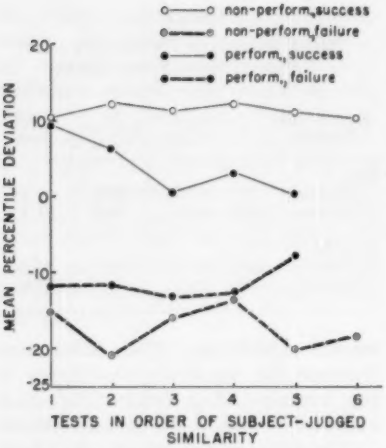


FIG. 1. Percentile deviations of each group's generalization test estimates from its vocabulary estimate.

Hypothesis 1, which predicted that positive and negative reinforcement have differential effects on subsequent expectancies, was confirmed. Figure 1 shows the mean percentile deviations of each of the group's estimates on the generalization tests from their own vocabulary estimates. Positive reinforcement raised while negative reinforcement lowered all subsequent expectancies. No temporal order is implied by the points along the abscissa as the presentation order of the tests differed for each *S* and was independent of his rankings of test similarity. An analysis of variance, found in Table 2, of the deviation scores shows that the expectancy differences between the success and failure groups were significant beyond the .001 level.

Hypothesis 2, which predicted that degree of generalization of expectancy changes decreases with increasing dissimilarity of the related tasks, was generally not confirmed. It was predicted that the slope of the success groups' curves would fall and the failure groups' curves would rise with increasing test dissimilarity. Figure 1 shows that the gradients for the success and failure groups which did not take the tests remain relatively flat. The gradients of the groups which took the tests give more but hardly conclusive support for the hypothesis. Although both the success and failure groups that took the tests show some decrease in generalization effect, only the success group's gradient shows a consistently decreasing generalization effect. The main analysis of variance results and the triple interaction effect between the performance, reinforcement, and task variables, significant at the .10 level, generally indicate that the increased dissimilarity of the tasks did not mediate any significant decrement in expectancy for the experimental conditions.

Hypothesis 3, which predicted that degree of generalization of expectancy changes decreases with increased *S* familiarity with the task, was gen-

TABLE 2

ANALYSIS OF VARIANCE OF DEVIATION SCORES BETWEEN VOCABULARY AND GENERALIZATION TEST[a] PERCENTAGE ESTIMATES

Source	*df*	*MS*	*F*
Between *Ss*	79		
Perf-NonPerf (PNP)	1	91.21	
Success-Failure (SF)	1	48,334.03	33.43**
PNP × SF	1	4,428.89	3.06*
Error	76	1,445.98	
Within *Ss*	320		
Tasks (T)	4	89.24	
T × PNP	4	172.17	1.18
T × SF	4	119.2	1.03
T × SF × PNP	4	235.26	2.04*
Error	304	114.87	
Total	399		

[a] Because the number of tasks for P and NP groups differed, the percentage estimates on the most dissimilar task of NP group are not included.

* $P < .10$.
** $P < .001$.

erally supported. Not only do the gradients of the groups that took the tests show some decrease in the extent of the generalization effect along the generalization continuum, but also the gradients show the effect of the initial reinforcement on subsequent generalization task estimates to be consistently smaller in magnitude. Figure 1 shows that the deviations from initial expectancy for every estimate of the groups which took the tests were less than the corresponding deviations of the groups which did not take the tests. A sign test analysis of this consistent directional effect was significant beyond the .001 level. The results suggest, therefore, that the increased familiarity with a situation reduces both the extent and magnitude of the generalization effect.

Less consistent results were obtained when each *S*'s test estimates were ordered on the basis of the group judged similarity of the generalization tests to the vocabulary. The test estimates were also analyzed with respect to the temporal presentation order for each *S*. No consistent results were discernible, thus suggesting that there was no tendency for the generalization effect to decrease with time irrespective of the nature of the tests.

The experimental conditions permitted an analysis of the generalization effect of positive and negative reinforcement in one test upon subquent test *performance* in other tests. The performance scores of each success and of each failure *S* on each of the five generalization tests were obtained irrespective of *S* judgments of similarity. Position effects were equated between the five tests since the temporal order of test presentation had been randomized. Table 3, which shows the differences in performance on each of the five tests between the success and failure groups, shows no significant differential effects. This result is contradictory to other studies which have shown that failure in one task may depress performance in other different tests (Lantz, 1945; Sears, 1937).

TABLE 3

DIFFERENCES IN GENERALIZATION TEST PERFORMANCE FOLLOWING SUCCESS AND FAILURE ON VOCABULARY

Test	Success		Failure		CR
	Mean	*SD*	Mean	*SD*	
Analogy Test	30.4[a]	9.33	30.8	9.78	.13
Par. Comp. Rec.	26.1	4.73	26.7	3.47	.4
Assump. Number	10.8	2.88	11.9	2.26	1.28
Series	17.8	4.4	18.8	3.56	.73
Dexterity	235.2	50.9	222.5	40.5	.71

[a] High score is best score for all tests except Dexterity Test.

In a postexperimental interview on the determinants of their expectancies, 25% of the *S*s spontaneously volunteered and 58.8% of the remaining *S*s admitted after direct questioning that the score obtained on the vocabulary test influenced their subsequent expectancies. The success *S*s' comments were usually affectively neutral while the failure *S*s' comments were marked by pervasive negative self references, e.g., "I did so lousily on the first test . . . made me more cautious . . . lost my courage after the first test . . . my confidence failed."

DISCUSSION

The experiment, using similar generalization tasks and procedures but a different method of data analysis, confirms and extends the findings of Rotter (1954) and his associates (Chance, 1952; Jessor, 1954) that positive and negative reinforcement of an expectancy affect

subsequent expectancies in a large number of increasingly dissimilar tasks. The results do not unequivocably confirm Jessor's (1954) finding that the generalization effect decreases with decreasing stimulus similarity or, as he defines the generalization continuum, decreasing goal-relatedness of the stimuli. The results of *Ss* who performed the tests show some, though certainly not marked, decrement in generalization effect, similar to that shown by Jessor's *Ss* who *also* performed on similar academic and physical skill tests. However, under the condition of not taking the generalization tests, increasing stimulus dissimilarity, or the goal-related values of academic and physical skill achievement, did not mediate any decrease in generalization effect within the range of tests used. Since many *Ss* appeared or reported to be under considerable tension during the experiment, it may be that *Ss* under high drive level, possibly induced by the experimental instructions, generalize more extensively than do *Ss* under reduced drive level. Rotter's thesis may need to be extended to take into account both the variables of test performance and drive level as additional determinants of the magnitude and extent of expectancy generalization.

The study gives no direct confirmation to Rotter's thesis that expectancy generalization is mediated by the goal relatedness of the stimuli since no independent measure of goal relatedness was obtained. Although scholastic aptitude tests may have a higher reinforcement value than physical skill tests for a college group, as the Rotter group claims, it seems more likely that objective physical similarity of the different tests in this experiment was the more important variable mediating the small decrement in generalization effect. When rating the tests many *Ss* classified them into paper and nonpaper tests and then proceeded to distinguish between the paper tests in terms of their use of words or numbers.

Summary

It was hypothesized that positive and negative reinforcement have differential effects on the generalization of expectancies, that the amount of generalization varies with stimulus similarity and with *S* familiarity with the task situation. Each *S* stated an expected percentile score on a vocabulary test, took the test and was given a prearranged score which was either below (failure) or above (success) his initial expectancy. Forty *Ss* then each estimated their performance on five or six randomly presented tests that varied in degree of similarity to the vocabulary. Another 40 *Ss* followed the same procedure but after each statement of expected performance on a particular test, performed the test. The major conclusions are: (*a*) positive and negative reinforcement have differential effects on the generalization of expectancies, (*b*) increased generalization task dissimilarity did not mediate any consistent decrement in generalization except for the performance success group, and (*c*) increased familiarity with a situation reduces the extent and magnitude of generalization. It was suggested that Rotter's theory of expectancy generalization needs to be revised to take into account determinants other than goal relatedness that affect the shape of the generalization gradient. Success and failure were also found to have no differential effect on subsequent scholastic aptitude or physical skill test performance.

References

Chance, J. E. Generalization of expectancies as a function of need relatedness. Unpublished doctoral dissertation, Ohio State Univer., 1952.

Jessor, R. The generalization of expectancies. *J. abnorm. soc. Psychol.*, 1954, **49**, 196–200.

Lantz, B. Some dynamic aspects of success and failure. *Psychol. Monogr.*, 1945, **59**, (1), (Whole No. 271).

Lewin, K., Dembo, T., Festinger, L., & Sears, P. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944, Ch. 10.

Rotter, J. B. Level of aspiration as a method of studying personality. I. A critical review of methodology. *Psychol. Rev.*, 1942, **49**, 463–474.

Rotter, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.

Sears, R. R. Initiation of the repression sequence by experience of failure. *J. exp. Psychol.*, 1937, **20**, 570–580.

(Received October 18, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

FRUSTRATION STIMULI IN DISCRIMINATION[1]

D. W. TYLER,[2] MELVIN H. MARX, AND GEORGE COLLIER

University of Missouri

Two frequently hypothesized effects of frustration are (*a*) an increase in drive level, and (*b*) production of new or unique stimuli arising as a consequence of this increased drive (Lawson & Marx, 1958b). Blocking, delaying, and failing to reward a previously rewarded response have been some of the operations presumed to produce frustration.

To examine the second of the two assumed consequences of frustration, Amsel and Ward (1954, Exp. I) trained rats to make one turn in a T maze when fed just prior to the choice point and to make an opposite turn when not fed. The authors assumed that in a sequence of rewards and nonrewards prior to the choice point, nonreward produces frustration, and that the stimuli arising from the frustration serve as cues to differential responding.

In this design, however, it is possible that *S*'s discriminative behavior could be based primarily upon some form of differential "peripheral stimulation" accompanying reward and nonreward.

The absence of the reward container on nonreward (frustration) trials is one possible source of uncontrolled peripheral stimulation in the Amsel and Ward studies. The only effort made to control this factor was to render the container relatively inconspicuous; it was small and was painted the same brightness as the goal box in which it was placed. In all likelihood, however, the container was still perceptible under this arrangement, since *S*s presumably encountered no difficulty in locating and consuming reward when it was present.

A second possible source of uncontrolled peripheral stimulation is the food-no food (F-Nf) contrast. Residual food particles and the consummatory responses associated with feeding might provide the basis for the discrimination. In order to control for mouth-cues in Exp. IV (Amsel & Ward, 1954), water, rather than food, served as reinforcement and *S*s were "pre-watered" in the starting box on all trials from the beginning of training. It was assumed that these conditions would minimize differences in mouth-cues at the choice point arising from reinforcement, as compared with nonreinforcement just prior to the choice point. Comparison of Exp. IV with Exp. I revealed no difference in rate of learning, suggesting that the critical cue in both cases was frustration stimuli.

The "pre-watering" technique, however, can be considered as only a partial control for this factor of differential mouth-cues. The time since the last reinforcement as well as the number of reinforcements (whether one or two) still varied under this procedure, depending upon whether *S* was rewarded or not rewarded in the stem goal box.

It would appear, therefore, that the Amsel and Ward experiments cannot be viewed as unequivocal support for the frustration stimulus notion because sources of differential stimulation other than possible frus-

[1] This investigation was supported by a research grant, M-817, from the National Institute of Mental Health, of the National Institutes of Health, Bethesda, Maryland.

[2] Now at Louisiana State University.

tration stimuli were confounded within the design. The present experiment attempted to eliminate this confounding in order to evaluate the independent contribution of presumed frustration stimuli. In general, this was accomplished in two ways. First, a food container with inaccessible reward was present in the goal box on nonrewarded trials throughout training.[3] Inaccessible reward should be at least as effective, if not more so, in producing frustration (Amsel & Hancock, 1957) but it should minimize differences in stimulation (visual, tactual, etc.) resulting from an empty goal box on nonrewarded trials. Second, a prefeeding procedure was instituted *after* *S*s had already learned a food-rewarded problem similar to that used by Amsel and Ward. The discriminative behavior of prefed *S*s was then compared with comparable controls to determine the *immediate* effect of prefeeding before new learning could take place.

Method

Subjects

The *S*s were 36 male albino rats obtained from the Psychology Department's colony. One *S* in Group F-Nf died during training. All *S*s were approximately 120 days old at the start of the experiment.

Apparatus

A runway and a T maze were used. The straight runway was 11 ft. 8 in. long, 3 in. wide, and 5 in. deep, with a 12-in. start box and two successive 18-in. goal boxes. The approach stem to Goal Box 1 was 6 ft. and that to Goal Box 2 was 20 in. A guillotine door separated the start box from the first approach stem; two guillotine doors formed Goal Box 1, and another separated the second approach stem from Goal Box 2. The start box and both approach stems were painted mid-gray while the two goal boxes were painted flat black.

The T maze was formed by detaching the second approach stem and goal box from the straight runway and replacing these with a shorter 9-in. stem and two 22-in. cross arms to form the T. A guillotine door 12 in. from the end of each arm formed a goal compartment. These were painted flat black while the remainder of the T was mid-gray. Both the straight runway and the T maze were covered by hardware cloth.

Three food cups were utilized. They were made from ½-in. high brass cylinders 2 in. in diameter. A small cup was drilled in the center of each. The inaccessible food cup covered with ¼-in. hardware cloth soldered to the top. Hardware cloth was also soldered to the top of the remaining two cups except for the depression in the center where it was cut away. All three cups were painted flat black to match the goal compartment in which they were placed.

Both the runway and the T maze rested upon a table 30 in. above the floor. Illumination was provided by a 200-w. incandescent light 7 ft. above and in the center of the apparatus.

Procedure

All *S*s were housed individually and were placed on a 22½-hr. deprivation schedule one week prior to the start of the experiment proper. Dry food was used. Water was present in the home cage at all times. The *S*s were run approximately 20 hr. hungry.

Accommodation.—On each of five days just prior to the start of runway training, *S*s were placed in the straight runway in groups of six for 5 min. and allowed to explore and eat Little Friskie pellets scattered throughout the runway.

Runway training.—Runway training began on the eighth day. The *S*s were randomly divided into three subgroups of 12 *S*s each. Half of the *S*s, six from each of the subgroups, were run on one day and the remaining half on the following day. Ten rotated (spaced) trials were given each day until a total of 60 trials had been administered.

On each trial *S* was required to traverse the runway from start-box to Goal Box 1 and then from Goal Box 1 to Goal Box 2. The reward given in each goal box was one Little Friskie pellet (approx. .1 gm.); the food cups were situated in the center on the runway floor at the ends of the goal boxes.

[3] The presence of the empty food cup on nonrewarded trials would have been perhaps an equally suitable control. The use of inaccessible reward precluded complete identity in food cups, of course, but permitted better control of such cues as visual stimuli arising from the food itself.

At the start of a given trial the entrance door to Goal Box 1 was in the raised position, the second door was in the lowered position, and the door to Goal Box 2 was raised. When *S* faced the starting-box door, this door was raised. As *S* entered Goal Box 1 the door behind it was lowered and *S* was retained from 5 to 10 sec. (The exact time varied randomly from trial to trial.) The exit door was then raised, permitting *S* to approach and enter Goal Box 2. Time required for *S* to traverse the first stem from the moment the starting-box door was raised until entry into Goal Box 1 was recorded manually by means of a stop watch.

Discrimination training.—Three groups of 12 *Ss* each were formed on the basis of performance during runway training. All *Ss* were given four straight runway trials using the procedure described above, followed by 10 discrimination trials on the T maze each day throughout all remaining portions of training.

Group F-Nf (Food-No food) was rewarded in Goal Box 1 on 5 of these 10 trials and not rewarded on the remaining 5 trials. On rewarded trials the goal box in the right arm of the T contained additional reward (one Little Friskie pellet) while the left arm contained neither food nor food cup. On nonrewarded trials the reverse was true. The task, therefore, was to learn to turn right when fed in the goal box in the approach stem and to turn left when not fed at this point.

Group F-If (Food-Inaccessible food) was treated exactly like Group F-Nf, except that on nonrewarded trials inaccessible food instead of no food was present in the approach stem.

Group If-Nf (Inaccessible food-No food) was never fed in the approach stem on any discrimination trial. However, on those trials corresponding to rewarded trials in the other two groups, inaccessible food was present in the approach stem and food was in the right arm of the T. On the remaining trials, neither food nor the food cup was present in the first goal box and food was placed in the left arm of the T.

On all trials *Ss* were retained in the stem goal box from 5 to 10 sec.

The *Ss* were run in squads of six, with trials rotated. Half the *Ss* from each of the three groups were run on a given day. The reward-nonreward pattern in the stem goal box was: RNRNNRNRRNNRNRRNRNRN.

Prefeeding.—After completion of initial T-maze training (340 trials), each of the three groups was divided into two matched subgroups on the basis of performance during discrimination training. Forty additional trials were given in which the procedure was identical to that followed throughout discrimination training, with one exception. One subgroup from each of the three major groups was randomly designated as the prefeeding subgroup while the remaining *Ss* served as controls. The prefeeding procedure consisted of presenting one Little Friskie pellet to *S* on each trial immediately prior to placing it in the start box. No *S* rejected the reward on any trial.

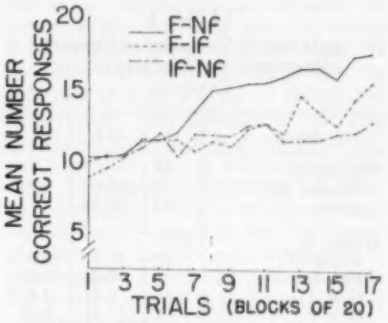


FIG. 1. Mean number of correct responses for the three groups as a function of training trials.

RESULTS

Discrimination training.—Figure 1 depicts the performance of the three groups during discrimination training. It shows the mean number of correct responses in blocks of 20 trials for each of the three groups. Table 1 presents an analysis of variance of the number of correct responses for the three groups, right-left position, and 17 blocks of 20 trials. To equate *Ns* in the three groups, one *S* selected randomly in Group F-If and one in Group If-Nf were omitted from this analysis. Groups, Trials, and the G × T interaction were significant.

Separate *t* tests, using the error term from the between analysis, revealed that Group F-Nf differed significantly from the other two groups, while Groups F-If and If-Nf

TABLE 1

ANALYSIS OF VARIANCE OF CORRECT RESPONSES FOR DISCRIMINATION TRAINING

Source	*df*	*MS*	*F*
Between *Ss*	32		
Groups (g)	2	188.07	7.95*
$error_b$	30	23.65	
Within *Ss*	1089		
Trials (T)	16	41.01	15.96*
Position (P)	1	0.47	0.02
G × T	32	6.75	2.63*
G × P	2	17.41	0.66
T × P	16	16.42	1.31
G × T × P	32	11.49	0.91
$error_w$	990		
e_1	480	2.57	
e_2	30	26.19	
e_3	480	12.57	

* $P > .01$.

did not differ significantly from each other. The performance of Group If-Nf was compared with the expected value ($M = 10$) and was found to be significantly greater ($t = 3.06$, $df = 30$, $P < .01$). The significant T × G interaction indicates that the three groups attained their respective levels at different points in training.

In the Amsel and Ward studies *Ss* made more responses early in training to the nonreward side of the T than to the reward side. The fact that neither the Position variable nor any interaction involving Position was significant in the analysis in Table 1 indicates that, considering all training trials together, *Ss* in the present experiment made about equal numbers of right and left turns in the T. However, an examination of the training data over the first 60 trials shows that there was a significant tendency to prefer the left side (side rewarded following nonreward) of the T early in training. This tendency was greatest in Group F-Nf and least in Group F-If.

Prefeeding.—The mean number of correct responses during prefeeding training for prefed and nonprefed (control) subgroups in each of the three major groups is presented in Table 2. Comparable means for the same *Ss* on the 20 trials preceding prefeeding training are also shown. To equate the number of *Ss* in the six subgroups, one additional *S*, selected randomly, in Group F-If and one in Group If-Nf were discarded. Consequently, each subgroup mean is based upon the performance of five *Ss*. An analysis of variance similar to that presented in Table 1 on Groups, Prefeeding, Position, and Blocks of trials showed the main effects of Groups and the Prefeeding × Blocks of trials interaction to be significant. The blocks of trials were the block of 20 trials preceding the introduction of prefeeding and the two

TABLE 2

MEAN NUMBER OF CORRECT CHOICES IN A BLOCK OF 20 TRIALS FOR LAST 20 DISCRIMINATION TRIALS AND THE TWO BLOCKS OF 20 PREFEEDING TRIALS

Group	Pretest (Last 20 Trials during Original Training)		Test 1 (Prefeeding, First 20 Trials)		Test 2 (Prefeeding, Second 20 Trials)	
	Fed	Control	Fed	Control	Fed	Control
F-Nf	17.6	17.0	16.8	17.0	12.8	18.6
F-If	15.2	15.4	12.6	14.0	12.4	16.4
If-Nf	13.8	11.2	12.4	11.6	12.2	11.4

Note.—The pooled *SD* for comparisons between groups is .56 and for (within *Ss*) comparisons between trials is .33.

following blocks of 20 trials. The values of the Blocks of trials × Prefeeding × Groups means are given in Table 2. From these values and the significant Blocks of trials × Prefeeding interaction it can be seen that the introduction of prefeeding resulted in an immediate decrement in correct choices. The first block of trials for Group F-Nf appears to be an exception since little change occurred.

Discussion

The superior performance of Group F-Nf relative to Group F-If during discrimination training as well as the above-chance performance of Group If-Nf during the same period indicates clearly that stimuli associated with the presence of the food cup or the sight and smell of food on part of the trials but not on others formed a significant source of differential "peripheral" stimulation in the present experiment. Also, the fact that prefed *Ss* made significantly more errors than nonprefed controls in the second portion of the study indicates that differences in cues resulting from alternately feeding and not feeding *Ss* prior to the choice point formed a second significant source of differential "peripheral" stimulation which must be taken into consideration. The only question to be resolved, therefore, is whether it is necessary to invoke the operation of hypothetical frustration stimuli, in addition to these other demonstrated sources, in order to give a reasonable account of the data.

So far as the results of discrimination training are concerned, it may be argued that *Ss* of Group F-Nf learned to turn left in the T whenever they approached the choice point (*a*) frustrated, (*b*) without the effects of prefeeding, and (*c*) without the food cup present in the stem goal box; and to turn right whenever they approached the choice point (*a*) *not* frustrated, (*b*) *with* the effects of prefeeding, and (*c*) *with* the food cup present in the stem goal box. By this reasoning, the inferior performance of Group F-If relative to that of Group F-Nf resulted from the fact that *Ss* in Group F-If had the benefit of only the first two of these sources of differential stimulation. The presence of the food cup on all trials in this group precluded the possibility of the food cup serving as a discriminative stimulus for these *Ss*. Similarly, Group If-Nf learned least rapidly because these *Ss* had the benefit of only the cup-no cup difference in stimulation. Group If-Nf *Ss* were never fed in the stem goal box and presumably were frustrated on all trials.[4]

While this line of reasoning appears to account for the results of discrimination training, there is one difficulty. The similarity in rate of learning in Groups F-If and If-Nf suggests that the frustration cue played a relatively unimportant role in the discrimination learned by Group F-If. That is, the food-no food cue plus the frustration-no frustration cue in Group F-If produced only slightly faster learning than the cup-no cup cue alone did in Group If-Nf. This occurred in spite of the demonstrated potency of the prefeeding.

It appears that a preferable account of these results is provided by consideration of the same peripheral cues without making reference to hypothetical frustration stimuli. According to this second interpretation, Group F-Nf learned most rapidly because *Ss* in this group had the benefit of the cup cue in conjunction with the food cue. The remaining two groups encountered a more difficult discrimination and therefore, learned more slowly because only a single source of differential stimulation (cup-no cup in Group If-Nf and food-no food in Group F-If) was available to these *Ss* at the time of choice.

[4] There is no direct evidence available, outside of face validity, that inaccessible food will produce "frustration," particularly in a sequence of presentations in which food is never consumed. As a matter of fact, there is experimental evidence that under such conditions inaccessible food acts as a reinforcer (Lawson & Marx, 1958b; Schlosberg & Pratt, 1956).

Besides being more parsimonious, there there are two reasons why an account in terms of peripheral cues alone appears preferable to one which invokes frustration stimuli as a supplementary source of differential stimulation.

First, the similarity in rate of learning in Groups F-If and If-Nf is more readily incorporated within this view. To do so, it is simply necessary to assume that the presence and absence of food in the mouth (Group F-IF) provided stimulation only slightly more discriminable than that resulting from the presence and absence of the food container (Group If-Nf). Both cues in combination (Group F-Nf), however, were significantly more discriminable than either cue alone.

Second, the performance of *S*s following prefeeding appears to fit more readily into a peripheral cues interpretation. The prefeeding operation was designed to minimize differences in mouth and associated cues but presumably did nothing to attenuate frustration. Yet, the absence of differential prefeeding cues appears to have resulted in almost total breakdown of the discrimination in Group F-If, suggesting that, even if present, frustration stimuli were of no consequence in directing these *S*s' choice behavior.

Prefeeding had little effect on the performance of *S*s in Group F-Nf over the first 20 trials, which is understandable since these *S*s still had the benefit of one peripheral cue (cup vs. no cup) throughout prefeeding. Why the prefeeding operation should suddenly produce a decrement in the second block of 20 trials is less easily understood.

The reasoning with respect to the frustration hypothesis in the present experiment is necessarily indirect, since no direct evidence for the production of frustration, such as increased vigor, is available. The magnitude of reinforcement is of the same order of magnitude as that reported by Amsel (1958). The recent experiment of Seward, Pereboom, Butler, & Jones, 1957 suggests that the evidence for frustration in this situation is an artifact of prefeeding. In the present experiment the reduction in correct choices in Group If-Nf, which did not have the differential effects of feeding, suggests that prefeeding not only lowers running speed but also disrupts discriminative behavior.

The question may be raised as to whether a test situation of this general kind can ever be designed which will provide unequivocal support for inferring the existence of frustration stimuli. With the rat as an experimental *S*, the answer appears to be no. When differential responding is attributed to the products of differential "central" states (e.g., frustration, drive, etc.), the experimental operations by which the central states, and therefore stimuli arising from these states, are manipulated, must not, in themselves, enter directly into the discrimination. Although far from perfect, the usual control for this factor is to separate experimental operations and choice behavior by long time intervals. In the case of drive discrimination studies, for example, a long time elapses between the choice in the experimental apparatus and the depriving operation. Thus the possibility of an interaction between choice and the depriving operations is virtually precluded.

It is because of the seemingly transitory nature of the effect being dealt with in studies of frustration, at least in the rat, that this important temporal separation between differential experimental operations and choice behavior does not appear to be possible. Therefore, what one investigator terms response to "central" (frustration) stimuli, another may with equal, if not greater, justification term response to "peripheral" (environmental) stimuli.

SUMMARY

The purpose of the experiment was to evaluate the role of assumed frustration stimuli in discrimination learning. Three groups of 12 rats were trained on discrimination problems in a T maze. Group F-Nf was required to turn right in the T whenever fed just prior to the choice point and to turn left when not fed; Group F-If was trained the same except that on nonrewarded trials inaccessible food was in the goal box prior

to the choice point; Group If-Nf was never fed in the stem goal box and was required to turn right when the stem goal box contained inaccessible food and to turn left when the stem goal box was empty. After 340 training trials, the groups were split into two matched subgroups, one of which was prefed in the start box on each of 40 additional trials, while the other continued on as in discrimination training.

Group F-Nf learned significantly more rapidly during discrimination training than either of the other two groups. Group F-If was superior to If-Nf during the same period but not significantly so. In the second portion of the study, prefed *Ss* made significantly more errors than nonprefed controls.

It was concluded that the results of the present study could be accounted for in conventional discrimination terms without recourse to the concept of frustration stimuli.

References

Amsel, A. Comment on "role of prefeeding in an apparent frustration effect." *J. exp. Psychol.*, 1958, **56**, 180–181.

Amsel, A., & Hancock, W. Motivational properties of frustration: III. Relation of frustration effect of antedating goal factors. *J. exp. Psychol.*, 1957, **53**, 126–131.

Amsel, A., & Roussel, J. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.

Amsel, A., & Ward, J. S. Motivational properties of frustration: II. Frustration drive stimulus and frustration reduction in selective learning. *J. exp. Psychol.*, 1954, **48**, 37–47.

Lawson, R., & Marx, M. H. Frustration: Theory and experiment. *Genet. Psychol. Monogr.*, 1958, **57**, 393–464. (a)

Lawson, R., & Marx, M. H. A comparison of some presumed frustrating and secondary-reinforcing operations. *J. comp. physiol. Psychol.*, 1958, **51**, 742–746. (b)

Schlosberg, H., & Pratt, C. H. The secondary reward value of inaccessible food for hungry and satiated rats. *J. comp. physiol. Psychol.*, 1956, **49**, 149–152.

Seward, J. P., Pereboom, A. C., Butler, B., & Jones, R. B. The role of prefeeding in an apparent frustration effect. *J. exp. Psychol.*, 1957, **54**, 445–450.

(Received November 3, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

THE EFFECT OF "SOCIAL" DISCRIMINATIVE CUES ON PROBABILITY LEARNING[1]

EDITH D. NEIMARK

New York University

AND

SEYMOUR ROSENBERG

University of Kansas

It has become increasingly common to analyze interpersonal phenomena in terms of learning theory. In learning theory terms, social situations may be characterized as those in which the behavior of one person, or persons, provides either reinforcing or discriminative stimuli, or both, for another person, or persons (Hall & Rosenberg, 1957; Keller & Schoenfeld, 1950; Rosenberg & Hall, 1958; Sidowski, Wyckoff, & Tabory, 1956; Skinner, 1953). The present study is concerned with effects upon individual behavior of "social" discriminative stimuli in a highly simplified laboratory situation.

The experimental task is the verbal conditioning situation extensively studied by a number of investigators (Estes & Straughan, 1954; Grant, Hake, & Hornseth, 1951; Humphreys, 1939; Neimark, 1956). In this situation one of two events, E_1 or E_2, occurs on each trial with relative frequencies of π and $(1 - \pi)$ respectively. The *S*'s task is to predict which of the two events will occur by making the appropriate response, A_1 or A_2. It has generally been found that as the number of trials, n, becomes very large, the probability of an A_1 response, P_{A1} (n), approaches π. This "probability matching" has been predicted from theoretical considerations (Bush & Mosteller, 1955; Estes & Straughan, 1954).

We have added to this situation a discriminative stimulus which is reliable (i.e., correctly predicts E_1) on a proportion ρ of the trials and is reversed on the remaining proportion of them $(1 - \rho)$. The discriminative stimulus is characterized to *S* as the choice of one of his fellow *S*s, although it is actually programmed by *E*. It will be noted that when $\rho = 1$ and when $\rho = 0$, the situation reduces to the imitation paradigm of Miller and Dollard (1941). In the present study five values of ρ and three values of π have been used to determine the effect of a "social" discriminative stimulus upon response in a verbal conditioning situation. An additional purpose of the study was to determine if the statistical learning theory model for verbal conditioning (Estes

[1] This research was done under ARDC Project No. 7737 in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use and disposal in whole and in part by or for the United States Government. This study has not received the review and approval normally given to this Center's reports. The views, conclusions, and recommendations expressed or implied herein do not necessarily reflect the official views or policies of the United States Air Force or of the Air Force Personnel and Training Research Center. A preliminary report of these findings was presented at the meetings of APA, Washington, D. C., September 1958.

The authors express their appreciation for the assistance of the following people: Earl Lingle, James McDade, Richard Hall, John Roller, and Roland Rosebrock.

& Straughan, 1954) applies to prediction and discrimination behavior in this "social" situation.

METHOD

Apparatus.[2]—The apparatus consisted of a control unit operated by *E* and three identical *S* units. The *S* unit consisted of two boxes stacked one upon the other. Each box contained two panel lamps, 1 in. in diameter, mounted 7 in. apart. A spring-loaded toggle switch was mounted between the lower pair of lights. The switch could be moved through an arc of 50° in either direction from the center; it automatically returned to the center position when released. All movements of the switch were automatically recorded by an Esterline-Angus operations recorder. The *S*s were concealed from each other by plywood screens.

Each trial began with the lighting of one of the top pair of white lights, s_1 or s_2. The light remained on for a period of 5 sec., during which time *S* was to move his switch. After *S* had responded, one of the lower pair of green lights, E_1 or E_2, came on for the remainder of the 5-sec. period. Neither green light appeared if *S* responded after the termination of s_1 or s_2. A trial sequence was started by *E* and automatically timed and operated during the trial. Stimulus onset followed *E*'s pressing of the start button by .5 sec. during which period relays clicked as they did when an *S* responded. Thus it was reasonable for *S*s to believe that another *S* was operating a switch during this interval, and that this, in turn, had produced s_1 or s_2. The intertrial interval varied between 2 and 6 sec. with a mode at about 4 sec. All three *S*s run at a given time received the same sequence of s_i and E_i lights.

Subjects.—The *S*s were 135 Air Force recruits who were run three at a time. In each group of three *S*s one had an AFQT score of 93 or above, the second had a score of 91–65, and the third had a score of 63–31. Approximate Stanford Binet equivalents of these scores are: 127 and above, 126–108, 107–94, respectively (Gragg & Douglass, 1956). Since AFQT category had no statistically significant effect in any analysis, it will not be discussed below.

Design.—A factorial design consisting of all possible combinations of five values of ρ (1.00, .75, .50, .25, 0) and three values of π (1.00, .75, .50) was used. Three replications of the 15-group design were run. There were three *S*s in a replication of each combination. In two of the replications, Social groups, *S* was told that the upper white lights indicated the choice of one of his fellow *S*s. The two replications of the Social groups differed from each other with respect to the random sequence of s_i and E_i events. In the third replication, Nonsocial groups, *S*s were simply told that the upper white lights were a signal to predict. The random sequence of lights for this group was the same as that for one replication of the social groups. Wherever Social and Nonsocial groups are compared hereafter, only the one Social group having the same sequence as the Nonsocial group will be involved.

Procedure.—Each of the three *S*s was seated in his booth and told that since different instructions might be given to each *S*, he was to put on ear phones to receive instructions. The same tape-recorded instructions were then given to all three *S*s.

"The box in front of you contains two white lights on top and two green lights below with a switch between them. Your job will be to predict which of the two green lights will come on. If you think the green light on the right will come on, push the switch to the right; if you think the green light on the left will come on, push the switch to the left. The object is to predict the green light correctly just as many times as you can."

"First one of the two white lights will come on. *These indicate the prediction of one of the other two men as to which of the green lights he thinks will come on. You may agree with him or disagree with him. In either case you are not to make your decision until after you have seen your partner's choice.* This is your signal to indicate your prediction by throwing the switch either to the right or to the left. You must always make your prediction after one of the white lights appears and before it goes off. Predictions made after the white light is off do not count. To make your prediction, throw the switch either to the right or the left and then return it to the center position. After you have made your prediction one of the green lights will go on to tell you whether you were correct *or your partner was correct or neither or both of you were correct.* Remember, the object is to predict correctly the side on which the green light appears just as many times as possible."

The instructions were then repeated, after which *S* was told he would have four practice

[2] The apparatus was designed and partially constructed by R. M. Patton, whose assistance in completion of the apparatus is gratefully acknowledged.

TABLE 1

ANALYSIS OF VARIANCE OF A_1 AND I RESPONSES ON THE LAST FOUR 10-TRIAL BLOCKS FOR EACH GROUP

(Only values of F significant at the .05 level or beyond are included; nonsignificant effects have been pooled.)

Source	*df*	Response Measure			
		A_1 Responses		I Responses	
		MS	*F*	*MS*	*F*
Between *Ss*	89	15.52		44.81	
Groups (S vs NS)	1	28.34	5.02*	.18	
ρ	4	10.52		764.94	82.54***
$\rho \times$ Group	4	3.40		23.76	2.56*
π	2	430.56	76.34***	8.56	
Pooled effects	18	5.45		14.46	
Error	60	5.64		9.27	
Within *Ss*	270	.74		.98	
Trial $\times \rho$	12	1.43	2.17*	2.19	2.69**
Tr. $\times \rho \times \pi$	24	.82		1.82	2.24**
Pooled effects	54	.81		.92	
Error	180	.66		.81	
Total	359	4.41		11.85	

* P = .05.
** P = .01.
*** P = .001.

trials, and was asked to remove his ear phones. All *Ss* received the same instructions with the exception that sections in italics were omitted for the Nonsocial groups.

RESULTS

Two response measures were used: (*a*) the proportion of A_1 responses (predictions that E_1 would occur) in a block of 10 trials; (*b*) proportion of responses in a 10-trial block which were identical with "partner's" choice (s_i). The second class of responses will be called I responses hereafter. Although group variances were not generally homogeneous, and would not be expected to be so, analysis of variance was used for all comparisons which follow despite failure to fulfill the assumption of homogeneity of variance (Lindquist, 1953).

Proportion of A_1 responses.—The proportion of A_1 responses in successive 10-trial blocks for all *Ss* having the same π value are plotted for Nonsocial groups in Fig. 1 and for Social groups in Fig. 2. Frequencies of A_1 responding on the last four 10-trial blocks were compared by analysis of variance (see Table 1). As would be expected, π seems to be a determinant of final level of A_1 responding. There is also a significant difference between final levels for Social groups and Nonsocial groups. In general, Social groups seem to make fewer A_1 responses. Although the difference in final level is most marked in the case of the groups for which $\pi = .75$, the $\pi \times$ Group interaction falls short of significance at the .05 level.

In order to examine differences among Social and Nonsocial groups, a separate analysis of variance was done for the final level of A_1 responding for each group. In the case of the Nonsocial groups, the effect of ρ is significant ($F = 4.35$, 4 and 30 *df*). For the Social groups, only the Trial $\times \rho$ and Trial $\times \rho \times \pi$ effects

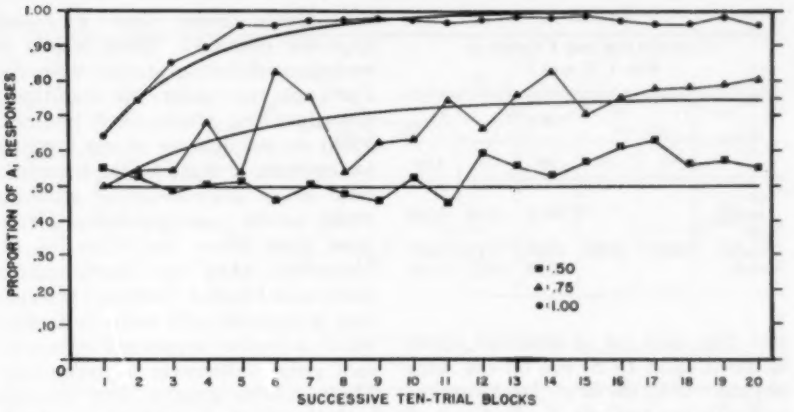


FIG. 1. Proportion of A_1 responses on successive 10-trial blocks for Nonsocial groups differing with respect to π. $N = 15$ *Ss*. Theoretical curves through the observed points have been derived from: $\bar{P}(m) = \pi - (\pi - \bar{P}(1))(1 - \theta)^{10(m-1)}$ where $\pi = .50, .75, 1.00$.

are significant ($F = 2.74$, 12 and 90 *df*; $F = 1.76$, 24 and 90 *df*). Why signal reliability, ρ, should have a greater effect for the Nonsocial groups, and why $\rho \times$ Group interaction should not be significant is not at all clear. Most probably the much smaller error variance for the Nonsocial groups is responsible for the significant ρ effect.

Theoretical curves through the observed points of Fig. 1 and 2 have been derived by the procedure suggested by Estes and Straughan (1954). Since the N for any combination of ρ and π is small, group curves were pooled across ρ (despite the questionable legitimacy of this procedure in view of the significant ρ effect for the Nonsocial group). In view of the

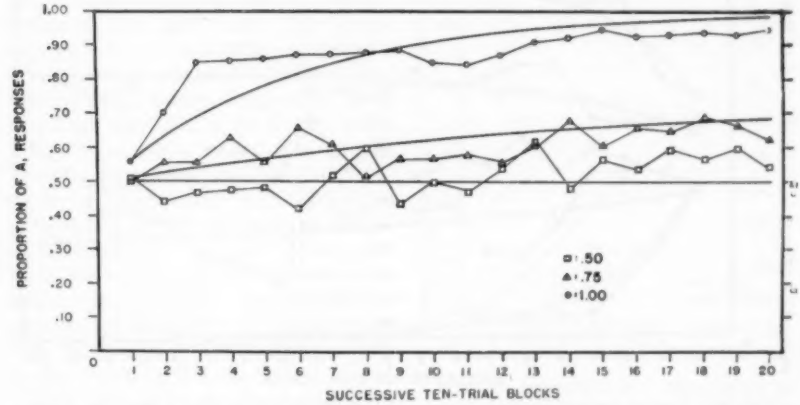


FIG. 2. Proportion of A_1 responses on successive 10-trial blocks for Social groups differing with respect to π. $N = 15$ *Ss*. Equation of the theoretical curves is the same as for Fig. 1.

TABLE 2

θ VALUES FOR THE CURVES IN FIG. 1, 2, AND 3

Group	π or ρ				
	0	.25	.50	.75	1.00
Non-social			.0078	.0236	.0329
Non-social	.0381	.0080	.0918	.0216	.0447
Social			.0024	.0071	.0170

fact that each set of observed points is based upon 15 *S*s run on the same sequence of E_1 the fit of the theoretical curves is not too bad.

Although the *S* population and experimental procedure of the present study differ from those employed in more traditional verbal conditioning studies (e.g., Estes & Straughan, 1954; Neimark, 1956) it is of interest to compare the present θ values with those obtained in the above studies (see Table 2). In general, the present θ values also seem to increase as a function of $\pi - P(1)$ but they are consistently lower than θ values typically obtained. Since results of verbal conditioning studies with Air Force *S*s run under the traditional procedure (e.g., Neimark & Shuford, 1959) do not indicate slower learning for recruits, it must be the introduction of a discriminative stimulus which serves to retard learning rate (and thus lower the value of θ). Moreover, when the discriminative stimulus is labelled "partner" learning rate is retarded still more: a finding which is further supported by significant group differences in asymptote. These results suggest that existing models of verbal conditioning (Bush & Mosteller, 1955; Estes & Straughan, 1954) apply in the present situation but that learning rate is a function of experimental conditions.

Proportion of I responses.—The proportion of responses matched with "partner," I responses, on successive 10-trial blocks for all subgroups with the same ρ value is plotted for Nonsocial groups in Fig. 3 and for Social groups in Fig. 4. Frequencies of I

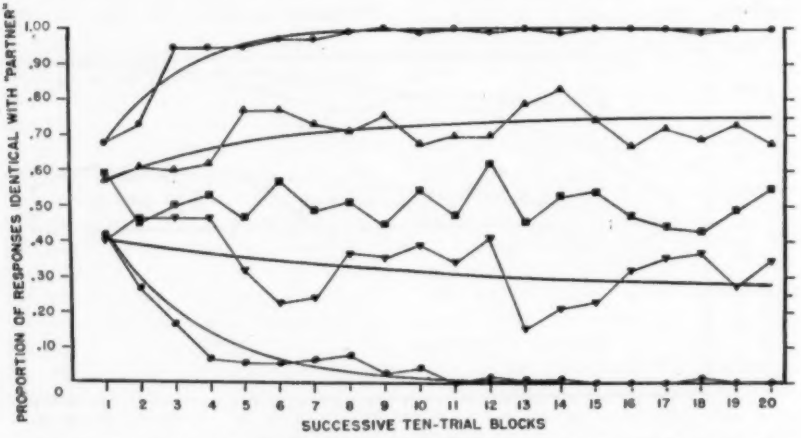


FIG. 3. Proportion of I responses on successive 10-trial blocks for Nonsocial groups differing with respect to ρ. $N = 9$. Theoretical curves through the observed points have been derived from: $\bar{P}_1(m) = \rho - (\rho - \bar{P}_1(1))(1 - \theta)^{10(m-1)}$ where $\rho = 0, .25, .50, .75, 1.00$.

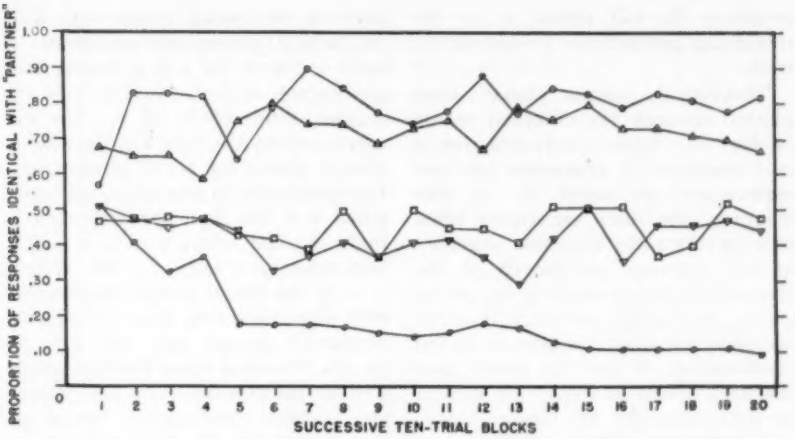


FIG. 4. Proportion of I responses on successive 10-trial blocks for the Social groups. For each point $N = 9$ *Ss*.

responses on the last four 10-trial blocks were compared by analysis of variance (see Table 1). The F for ρ effects is highly significant while that for π is not. Comparison of Social and Nonsocial groups does not yield a significant F ratio for between-group differences but does yield a significant F for Group $\times$ ρ interaction. An orthogonal polynomial analysis (Grant, 1956) of Group $\times$ ρ interaction shows that only the linear component is statistically significant ($F = 8.36$, 1 and 60 *df*). The function relating ρ and proportion of I responses appears to be linear, with a shallower slope in the case of the Social groups (see Fig. 5), i.e., the closer ρ approaches 1.0 or 0, the greater the difference between Social and Nonsocial groups. There are also two significant trial interactions which are probably attributable to differential variability in I responding as a function of differential variability in the sequence of s_1 and E_1.

The derivation of theoretical predictions concerning I responses proceeds in a manner completely analogous to that for A_1 responses (Bush & Mosteller, 1955; Estes & Straughan, 1954). In this instance, however, the two response classes are not A_1 and A_2 but I (A_1 responses following s_1, and A_2 responses following s_2) and NI (A_2 responses following s_1, and A_1 responses following s_2). The E_i appropriate to s_i will occur on a proportion ρ of the trials while the inap-

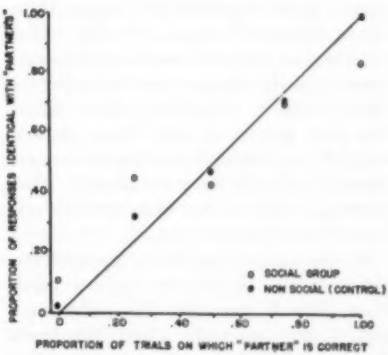


FIG. 5. Proportion of I responses on the last 40 trials as a function of ρ. The straight line is the value which would be predicted if complete probability matching obtained.

propriate E_i will follow s_i on the remaining proportion, $1 - \rho$, of the trials.

Theoretical curves have been plotted through the observed points of Fig. 3. Again, pooling across π is a questionable procedure but one necessitated by small N. In this instance, the data are based upon nine *S*s run under the same sequence of s_iE_i pairings, so the fit of the theoretical curves seems to be pretty good. As further evidence that responding to cue is analogous to verbal conditioning, it may be noted that the data points of Fig. 3 are similar to data obtained by Grant, Hake, and Hornseth (1951) who used π values identical with the ρ values of the present study.

Once again, the obtained θ values are somewhat lower than those typically reported, although they are similar to those obtained in this study for comparable π values (see Table 2). With the exception of the groups for which $\rho = .50$, there is also a suggestion that θ is a function of $\rho - \hat{P}_1(1)$. It will be noted, however, that the values are not symmetric, i.e., θ values obtained when $\rho > .50$ seem to be consistently larger than those obtained when $\rho < .50$. Discrimination learning seems to proceed more rapidly when a cue has a better than chance reliability than when the cue has a poorer than chance reliability, although the two situations should logically be equivalent. This question will be treated in more detail in the following section.

In the case of the Social group data (Fig. 4) no theoretical curves have been fitted because the observed points are generally quite far from their predicted asymptotes; for the groups with $\rho = .25$ one might even conclude that no learning has occurred. Although analysis of variance of the Social group data fails to yield a statistically significant F ratio for π or for $\rho \times \pi$ interaction, the curves of individual groups do suggest some effect of π. For example, when $\rho = 0$ or 1.0, learning is always slower for Social groups than for comparable Nonsocial groups (even when $\pi = 1.0$) but the difference is most marked when $\rho = 0$, $\pi = .75$, and when $\rho = 1.0$, $\pi = .50$. When $\rho = .25$ the Social groups show markedly slower learning than comparable Nonsocial groups only for $\pi = .50$ or .75. None of these findings augur well for the extension of a one parameter verbal conditioning model to the description of I responding in Social groups.

Comparison of I responding in groups with complementary ρ values.—In order to test the assumption that learning rate is the same for groups which are logically equivalent, curves for groups with complementary ρ values were compared by analysis of variance. The results are summarized in Table 3. Curves for groups with a completely reliable cue, $\rho = 1.0$, differ significantly from curves for groups with a completely unreliable cue, $\rho = 0$, in slope but not in over-all level. This supports the suggestion of faster learning for reliable cues than for unreliable ones. When the cue is partially reliable or partially unreliable, $\rho = .75$ and $\rho = .25$, curves of I responding differ in level but not in slope; level is higher for groups with $\rho = .75$. Bush and Mosteller (1955) in their analysis of data collected by Schwartz in a similar experiment found that events which increase imitation had a smaller effect than events which inhibit imitation. Their finding implies that learning of nonimitation should proceed more rapidly than learning of imitation—a prediction which is not

TABLE 3

ANALYSIS OF VARIANCE OF I RESPONSE CURVES FOR GROUPS WITH COMPLEMENTARY ρ VALUES: $\rho = 0$ OR 1.0; $\rho = .25$ OR .75

(Only values of F significant at the .05 level or beyond are included; nonsignificant effects have been pooled.)

Source	*df*	Comparison			
		0 vs. 1.0		.25 vs. .75	
		MS	*F*	*MS*	*F*
Between *Ss*	35	70.03		46.64	
Group (G)	1	358.42	6.87*	26.45	
ρ	1	.80		154.93	4.27*
π	2	47.51		204.80	5.64**
G × ρ × π	2	242.20	4.64*	22.58	
Pooled	5	52.24		25.15	
Error	24	52.14		36.28	
Within *Ss*	684	2.27		2.25	
Trial blocks (T)	19	27.06	23.33***	10.93	5.84***
T × π	38	.87		5.30	2.83***
T × ρ	19	3.36	2.90**	.81	
T × G	19	1.62		2.37	
T × ρ × G	19	2.01	1.73*	2.06	
T × π × G	38	4.16	3.59**	1.68	
T × ρ × π	38	2.72	2.34**	1.90	
T × ρ × π × G	38	2.25	1.94**	1.13	
Error	456	1.16		1.87	
Total	719	5.57		4.41	

* $P = .05$.
** $P = .01$.
*** $P = .001$.

supported by the results of this experiment.

The F ratios of Table 3 further suggest that when $\rho = 0$ or 1, curves of I responding for Social groups differ from curves for Nonsocial groups in level but not slope; interactions of Group × ρ × π, on the other hand, affect both level and slope of the curves. When $\rho = .25$ or .75 the social or nonsocial connotation of the cue has no effect at all. In this instance, sequence predictability, π, has a significant effect upon both slope and level of curves of I responding. These differences as a function of ρ are compatible with the data of Fig. 5.

Although the F ratios of Table 3 are anything but consistent and unequivocal they do seem to support the conclusions that: (*a*) reliable cues do not have the same effect upon learning as logically equivalent unreliable cues; (*b*) I responding is not independent of sequence predictability.

Generality across sequences.—The preceding comparisons have made use of only one half the *Ss* of the Social groups, i.e., three *Ss* in each subgroup. A replication of the Social group was run with a different random sequence of s_i and E_i in order to extend the generality of the findings for the Social groups across specific sequences. When frequency of A_1 responding is taken as a response measure, neither sequence effects nor the interaction of sequence with other variables yields an F ratio significant at the .05 level. This is also true of I responding as a response measure.

In the case of I responding, however, the F ratio for $\rho \times \pi$ interaction is significant at the .05 level ($F = 2.41$, 8 and 60 df). Although this interaction was not significant in the comparison of Social and Nonsocial groups, its effect was suggested by a comparison of group curves. This finding, then, further substantiates the conclusion that imitation learning in Social groups is a function of sequence predictability as well as cue reliability.

Discussion

The results of the present study show that existing formulations of verbal conditioning are applicable when a discriminative cue is introduced into the traditional verbal conditioning situation. The addition of a discriminative cue does, however, retard the rate of verbal conditioning (A_1 responding) and the retardation is even greater when the cue is given a social connotation.

That imitation, or nonimitation, can be learned and that the form of the learning curve is similar to that obtained in other situations was demonstrated by Miller and Dollard (1941). Our additional finding that imitation can be learned under conditions of partial as well as of continuous reinforcement extends the generality of their findings.

The extension of existing formulations of verbal conditioning to imitation learning does not provide an accurate description of group learning curves for the Social groups. Moreover, suggestions from the data of the present study that sequence predictability affects both slope and asymptote of the group curves would seem to suggest that no single-parameter model will provide an accurate description of imitation learning with "social" discriminative stimuli. Our findings that (*a*) Social groups tend to fall short of predicted asymptotes, (*b*) that differences between observed and predicted final levels increase as ρ approaches 0 or 1, and (*c*) that imitation learning proceeds more rapidly then nonimitation learning are all supported by results of a similar experiment by Kanareff and Lanzetta (1958).

Why learning should be consistently slower for Social groups remains a problem. One possible answer to this problem derives from the stimulus sampling concept of statistical learning theory (Estes & Burke, 1953). We assume that in any learning situation, but more especially in imitation learning with "social" discriminative stimuli, many of the stimulus elements are *S*-generated and based upon past experience in similar situations. For example, most *S*s probably have little past experience with consistently reliable or consistently unreliable social cues (facial expression, voicing of opinion, etc.) whereas neutral cues are more likely to be consistent (traffic lights, signs, printed schedules, etc.). All *S*s begin with attempts to figure out the experimental situation, but *S*s in the Social groups begin, in addition, with hypotheses of human fallibility and tendencies to wonder what the "partner" is up to. Thus, there should be a greater variety of associations and hypotheses in the case of the Social groups, i.e., in statistical learning theory terms, there should be a greater variety of stimulus elements, any one of which has a low θ_i weighting. On the basis of these assumptions, one would predict lower $\hat{\theta}$ values and, hence, slower learning in Social groups than in comparable Nonsocial ones. The rapid initial rise and slow approach to asymptote in I response curves of the Social groups are compatible with this analysis.

Summary

The effect upon verbal conditioning of the introduction of a probabilistic cue with a social connotation was studied by means of a factorial design comprising three values of event probability (E_1) and five values of cue reliability. One hundred and thirty-five *S*s received 200 trials in a modified verbal-conditioning situation. Two-thirds of the *S*s (Social group) were told that the cue represented choice of a partner; the rest (Nonsocial group) were not.

Response measures and results were: (*a*)

proportion of E_1 predictions was a linear function of E_1 frequency; (*b*) proportion of agreement with cue was a linear function of cue reliability. It was also affected by E_1 frequency. For both measures Social groups showed slower learning than Nonsocial groups.

Theoretical curves derived from an extension of statistical learning theory were fitted to the data. In the case of E_1 predictions, the theoretical curves provided a fairly accurate description of the data when extremely small θ values were used. The model was inadequate for description of agreement with cue in the case of the Social groups.

REFERENCES

BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning.* New York: Wiley, 1955.

ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.

ESTES, W. K., & STRAUGHAN, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

GRAGG, D. B., & DOUGLASS, H. J. Conversion tables for selected airman classification battery scores and comparable scores on other selected service and civilian tests. *Lackland AFB, Tex.: USAF Pers. Train. Res. Cent. Dev. Rep.*, 1956, No. AFPTRC-TN-56-89.

GRANT, D. A. Analysis of variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

GRANT, D. A., HAKE, H. W., & HORNSETH, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.

HALL, R. L., & ROSENBERG, S. An S-R analysis of social interaction. *Amer. Psychologist*, 1957, **12**, 370. (Abstract)

HUMPHREYS, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.

KANAREFF, V. T. & LANZETTA, J. T. The acquisition of imitative and opposition responses under two conditions of instruction-induced set. *J. exp. Psychol.*, 1958, **56**, 516–528.

KELLER, F. A., & SCHOENFELD, W. N. *Principles of psychology.* New York: Appleton-Century-Crofts, 1950, 352–397.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Miffin, 1953.

MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univer. Press, 1941.

NEIMARK, E. D. Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 209–220.

NEIMARK, E. D., & SHUFORD, E. H. Comparison of predictions and estimates in a probability learning situation. *J. exp. Psychol.*, 1959, **57**, 294–298.

ROSENBERG, S., & HALL, R. L. The effects of different social feedback conditions upon performance in dyadic teams. *J. abnorm. soc. Psychol.*, 1958, **58**, 271–277.

SIDOWSKI, J. B., WYCKOFF, L. B., & TABORY, L. The influence of reinforcement and punishment in a minimal social situation. *J. abnorm. soc. Psychol.*, 1956, **52**, 115–119.

SKINNER, B. F. *Science and human behavior.* New York: Macmillan, 1953. Pp. 297–312.

(Received November 10, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

THE EFFECT OF SHOCK ON THE RAT'S CHOICE OF A PATH TO FOOD[1]

THOMAS F. LOHR

Muhlenberg College

In certain situations shock appears to play a role quite opposite to its usual one. One of the more notable of these situations is that first described by Muenzinger (1934) in which rats shocked for making correct choices in a discrimination task learned more quickly than rats which were not shocked.

The hypothesis presented here stems from the observation that "shock for correct choice" means that shock is invariably followed by food reinforcement. The law of secondary or conditioned reinforcement tells us that any stimulus repeatedly followed by reinforcement acquires certain reinforcing properties of its own. In the absence of a reason for asserting that punishing stimuli must be an exception to the rule, we should *expect* that the aversive properties of shock repeatedly followed by food will be reduced and perhaps reversed in the course of a large number of trials.

There is, in fact, evidence from other situations that the aversive properties of the shock actually do undergo such changes when shock is repeatedly followed by food. Pavlov (1927, p. 30) reports that shock and other painful stimuli used as conditioned stimuli to salivation no longer elicit emotional response or the withdrawal reflex. Also, Masserman (1946) reports that some cats will "seek out" shock and air blasts which have been experienced in connection with the obtaining of food. It is well known that the sound of a food magazine which may originally startle the rat becomes an effective stimulus to increasing the speed with which bar pressing is acquired.

The experiments described below represent an attempt to test the hypothesis sketched above by introducing the shock-followed-by-food sequence into a new situation, one in which rats are shocked in one of two alleys *either of which* leads to food.

Experiment I

The first experiment was designed to detect what kind of force, aversive or reinforcing, was being exerted by the shock at any instant before, during, and after a certain number of shock-food combinations. The situation was arranged so that a rat had a choice of either a right or a left alley, both alleys being otherwise identical and both leading to the same goal box. Consequently, when shock was introduced into one of these alleys after the rat had established some pattern of responding, the effect of the shock could be continuously observed in terms of the changing ratio of right to left choices. Since all rats run in this apparatus without shock developed definite right- or left-going habits, it was possible to induce them to cross the shock grid a number of times without forcing by introducing shock onto the preferred[2] side after the habit was well established.

[1] These experiments represent a portion of a dissertation submitted to the Department of Psychology of Harvard University in partial fulfillment of the requirements for the Ph.D. degree.

[2] The term "preferred side" is to be understood as meaning the turn most frequently selected.

Method

Subjects.—Twenty-four brown rats of the Lashley strain between 100 and 150 days of age at the beginning of trials were used. Litter mates were distributed throughout the groups. All were handled daily for 10 days before the beginning of preliminary trials.

Apparatus.—The apparatus consisted of two alleys branching off from the vestibule of a starting box and converging again in the vestibule of a goal box. The door of the 8 × 8-in. starting box was the full width of the starting box and was hinged at the top. It opened inward permitting *S* to advance into the 5 × 8-in. vestibule. From this vestibule the alleys branched off to the right and left at an angle of 45°. They ran 4 in. in this direction, then turned and ran parallel to each other for 5 in. before turning back toward each other and converging again in another 5 × 8-in. vestibule. This led through a similar door into another 8 × 8-in. box. Since the apparatus was symmetrical from front to back as well as laterally, *S*s could be run back and forth without being removed between trials. The alleys were two inches wide except where they ran parallel to each other. Here they were 3 in. wide, each containing a 3 × 5-in. shock grid. Alleys were of wood 3 in. high and covered on top with wire screen.

The entire device was operated mechanically. An electric motor opened the doors at a signal from a timing device and closed the doors when *S* touched the food tray in the opposite goal box. The closing door started the timer and another device which dropped about ½ gm. of Gaines granulated dog food into the tray at the opposite end in preparation for the next trial.

The grids consisted of 24 copper strips glued to composition board. The potential supplied to the grid originated in a 66-v. AC transformer and was controlled by sets of fixed resistances which could be placed in series with the grid. For shock designated as "medium" the fixed resistance was 8,200 K., giving an estimated effective current through *S* of .07 ma. For shock designated as "strong" the fixed resistance was 22,000 K., giving an approximate effective current of .24 ma.

Procedure.—All *S*s were given preliminary training trials in which they were forced to the left and to the right in ABBA order for 10 trials on each of two days. All *S*s were then given 10 free-choice trials a day under 23-hr. food deprivation on consecutive days until the end of the experiment. They were run back and forth without being removed between trials, the interval being 2½ min. from the time *S* reached the tray until the doors opened for the next trial.

The 24 *S*s were divided into three groups of 8 each. Group C was allowed to run 500 trials and Group B 250 trials before shock was introduced. Shock was always given in the alley corresponding to the turn for which *S* had developed a marked preference. The *S*s in Group A were given shock beginning with the first trial, half in the left alley and half in the right alley. Half of each of the three groups was given shock of medium intensity and the other half was given shock of strong intensity. The shock intensity and number of preshock trials used were based on what preliminary experimentation suggested would be most likely to produce the desired results.

Results

Figure 1 shows for each *S* the number of right and left turns taken in the 10 trials of successive daily sessions. The tendency of the performance line for all *S*s in preshock trials to fall to the base line indicates the establishment of marked preferences for one turn or another. The immediate effect of introducing shock into the preferred alley varies, but the eventual effect is to precipitate behavior into two distinct categories. At the end of the shock trials all *S*s are either avoiding the shock turn 100% of the time or are selecting the shock turn 100% of the time. The number of *S*s falling into one of these categories or the other seems to vary with the intensity of the shock and, to a lesser extent, with the number of preshock trials. In any case, all *S*s given strong shock after 250 nonshock trials avoid the previously preferred turn consistently, while all *S*s given medium shock after 500 nonshock trials select the shock turn consistently. The performance of Group A shows that either of these intensities is sufficient normally to result in complete avoidance.

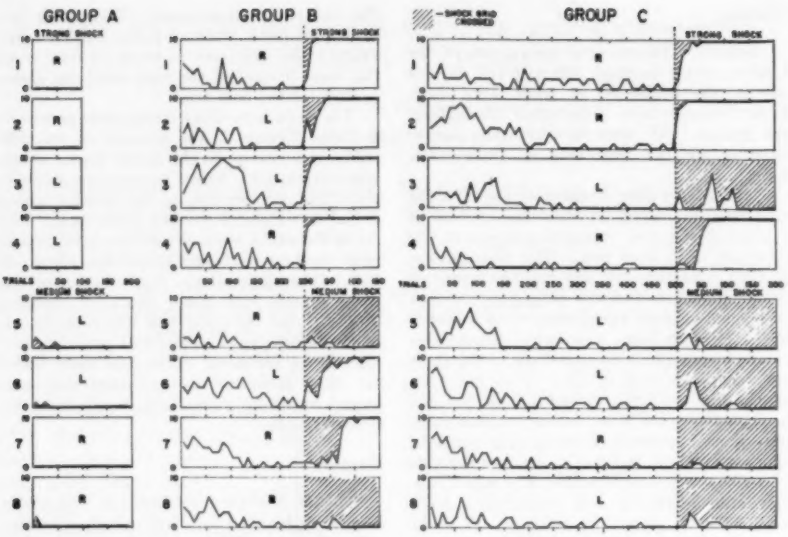


FIG. 1. The number of right and left turns made by each rat in successive daily sessions of 10 trials each in Exp. I. The number of turns to the nonpreferred side is below the line. Groups B and C were shocked at medium and strong intensity, respectively, on preferred turn beginning at the dotted line. Group A was shocked beginning on Trial 1.

Discussion

The acquired avoidance of the shock turn found in the majority of these rats reflects the usual disposition of animals to learn to avoid a punishing stimulus. What requires explanation is the fact that all *S*s in this situation do not do this. The concept of conditioned reinforcement provides a possible interpretation.

Looking at the typical behavior patterns of rats that did not avoid the shock turn, we see that the well-established preference becomes initially disrupted at the introduction of shock. This, too, reflects the aversive effect of a punishing stimulus. But in this case the avoidance tendency does not continue to increase. Instead, it appears to reverse as the *S* regains his previous level of 100% selection. This reversal must reflect a change in the aversive properties of the shock. In terms of secondary reinforcement, the change is due to the reinforcing properties accruing to the shock as *S*, under the impetus of his habit, accumulates more shock-followed-by-food experiences.

Interpreted in this way, the dynamics of the forces operating in this situation are analogous to those in rolling a ball toward the ridge of an incline. If the initial impetus is sufficient to carry the rat across the shock grid enough times to nullify the aversive effect, he will acquire what looks like a compulsion for taking unnecessary punishment. If, on the other hand, the impetus is not quite sufficient to carry past the critical point, he will show the normally expected complete avoidance. As we increase the intensity of the shock and decrease the initial strength of the habit, we increase the probability that any rat will fall into the avoidance category, and as we decrease the intensity of the shock and increase the habit strength through more preshock trials, we increase the probability that any rat will fall into the shock-preference category.

Experiment II

Although the first experiment shows that the aversive properties of shock will be modified, presumably as a result of being followed by food, it does not clearly show that shock will actually acquire positive reinforcing properties. The second experiment was an attempt to demonstrate such an effect by introducing a mild shock on one of the alleys from the first free-choice trial and gradually increasing the intensity.

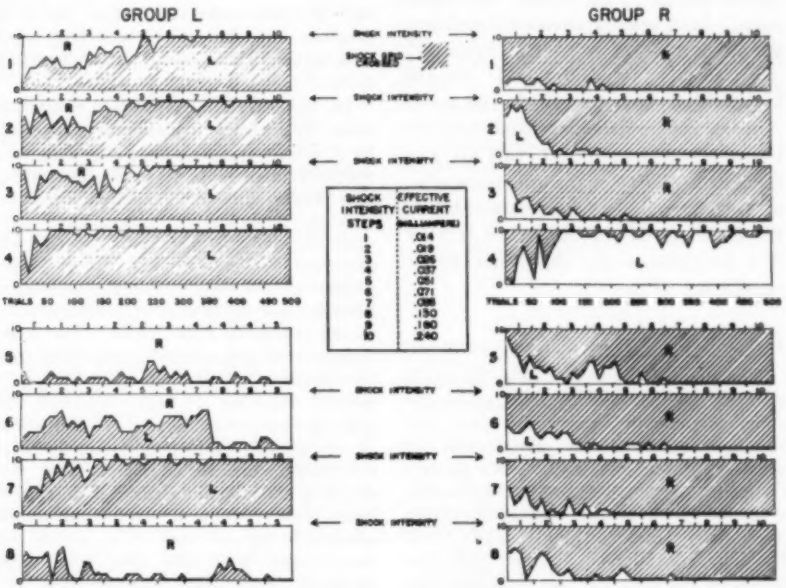
Method

In this experiment 16 rats were given preliminary training, as in Exp. I. The eight rats in Group L were given shock of intensity well below the avoidance threshold. They were shocked in the left alley beginning with the first free-choice trial. The eight rats in Group R were given the same shock in the right alley. The amount of shock on the grid for these two groups was increased one step at the end of every 50 trials, provided that *S* had crossed the grid at least 10 times in those 50 trials. The estimated effective current through *S* at each of the 10 steps from the most mild to the most severe is presented in the Table included in Fig. 2. It may be noted that Step 6 corresponds to what was called "medium" shock in Exp. I and Step 10 corresponds to "strong" shock.

The *S*s were run 10 trials a day for 50 consecutive days, in every other respect exactly as the *S*s in Exp. I.

Results

The performance of each *S* in Group L and Group R is shown in Fig. 2. As in Fig. 1, the performance line indicates the number of right and left turns taken in each session of 10 trials on consecutive days. How-


GROUP L
GROUP R
SHOCK INTENSITY
SHOCK INTENSITY
SHOCK INTENSITY

ever, in this figure left turns are found below the line and right turns above.

Two of the features to be pointed out in the performance of these *S*s are (*a*) the number of *S*s developing marked preferences for the shock turn within the first 200 trials, and (*b*) the lack of variability in choice by these *S*s in their last 200 trials. Here again, the patterns of performance divide into two distinct categories. Five of the eight *S*s in Group L and seven of the eight *S*s in Group R eventually selected the shock grid 100% of the time. These were all subject to regularly increasing intensities of shock. The remaining 4 of the 16 rats avoided the shock grid most of the time.

Perhaps more significant for these purposes is the consistency with which these 12 *S*s chose the shock turn. Not one of these 12 makes a single digression onto the nonshock alley in his last 100 trials. In the last half of their 500 trials the mean number of nonshock turns is 1.33. Using the *S*s in Group C of Exp. I as nonshock controls, the mean number of turns to the nonpreferred side in the last half of the 500 trials is 8.63. A t test shows the difference to be significant ($P < .01$).

Discussion

The surprising thing in these results is not that so many of the rats chose the shock turn with such consistency but that any of them did. We would not normally expect that a rat, when given a choice of a clear alley and one delivering a severe shock, would choose the shock alley all the time, especially when there has been no inducement other than the shock on the grid, for him to choose this alley rather than the other.

Again the interpretation offered is in terms of the acquired reinforcing properties of the shock. In the early stages of chance selection of alleys, the slight sensation of shock acquires reinforcing properties as a result of being followed by food. Consequently, *S* experiences a larger total of secondary reinforcement in the shock alley than he does in the nonshock alley. Once this process becomes operative, increases in shock intensity may be small enough that the conditioned reinforcing properties of the shock remain in excess of the aversive properties. The fact that these *S*s show a significantly greater consistency in making their preferred turn than do *S*s who are not shocked suggests further that the shock itself is making some positive contribution to the strength of the habit.

The fact that some *S*s do avoid the shock grid needs no special explanation. At any point in the process slight accidental factors might tip the delicate balance of positive and negative forces with a cumulative effect, in spite of the fact that the conditions most likely to produce the desired results had been determined as a result of preliminary tests.

In general, the secondary-reinforcement hypothesis, which was suggested by the improved learning found in rats shocked for correct choices in discrimination tasks, appears to meet with no serious obstacles in interpreting the data obtained in the two experiments presented here. On the other hand, neither of the two theories which have been applied to the discrimination experiments seems to apply here. The fear-reduction hypothesis originally presented by Mowrer (1940) interprets the improved learning by rats shocked for correct choices as being the consequence of an added reinforcing effect of fear reduction whenever *S* has crossed the shock grid. In the situation described here, fear reduction would be more effectively served were *S* to avoid the shock alley and take the other path to food. The other interpretation, offered by Muenzinger and others (Muenzinger & Fletcher, 1937; Muenzinger & Newcomb, 1936), to account for improved learning as a result of shock

for correct choices in a discrimination task is in terms of a certain emphasis lent to the discriminanda by the anticipation of shock. In the situations described here reinforcement does not depend upon discrimination at all since *S* gets food no matter which way he goes.

The secondary reinforcement interpretation of the discrimination experiments offers some other minor advantages. There has been some doubt as to whether any improvement in learning actually does result from shock for correct choices when *S* is not allowed to correct incorrect choices (Wischner, 1947). Certainly, the improvement is less marked in the noncorrective situation than it is in the corrective situation (Muenzinger & Powloski, 1951). This difference may be accounted for in part by the fact that in the corrective situation the rat must experience a shock-food sequence on every trial while in the noncorrective situation he does not. Both Wischner (1947) and Muenzinger (1948) have pointed out the rapid decline in errors when the shock-right *S*s finally begin to learn. This is to say that the positive effects of shock on learning take place mainly in the later trials. In terms of conditioned reinforcement such positive effects could only occur after *S*s had sufficient number of shock-food experiences and therefore must appear in the later trials. Muenzinger, Brown, Crow, and Powloski (1952) have shown that giving pretrials with shock increases the amount of improvement in shock-right rats. Prince's (1956) discovery that holding off shock until the twenty-fifth trial results in significantly faster learning than shocking them for correct responses from the beginning, may well be related to the observation in Exp. I that shock is less disruptive when introduced at a later stage of learning. It may be noted also that in Prince's experiment shock is not introduced at full strength in the beginning, but is built up from half strength in two steps in the course of 20 trials.

In conclusion, the secondary-reinforcement hypothesis appears to be useful in interpreting data relating to several situations in which shock appears to reverse its usual role. It has some additional strength in that it involves no assumption other than the well-established law of secondary reinforcement.

Summary

Two experiments were preformed in an apparatus which offered rats a choice of equivalent right and left pathways to a goal box containing food. When a strong shock was introduced into the preferred alley after 250 nonshock trials, all four *S*s avoided the shock turn consistently, but when a shock of lesser intensity, but still well above the normal avoidance threshold, was introduced into the preferred alley after 500 preshock trials, all four *S*s eventually chose the shock turn consistently. With strong shock after 500 trials or medium shock after 250 trials some *S*s fell into the avoidance category and the rest fell into the preference category. In Exp. II, shock of intensity well below the avoidance threshold was introduced into one alley, eight rats receiving shock for right turns and eight rats receiving shock for left turns. The shock was gradually increased to severe intensity in the course of 500 trials. Of the 16 *S*s, 12 chose the shock turn 100% of the time in the last 100 trials. An attempt is made to interpret these data in terms of secondary reinforcing properties acquired by the shock as a result of having been repeatedly followed by food. This hypothesis is also applied to the interpretation of other results in which the aversive effects of punishing stimuli seem to be reversed.

References

Masserman, J. H. *Dynamic psychiatry.* Philadelphia: Saunders, 1946.

Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.

Muenzinger, K. F. Motivation in learning: I. Electric shock for correct responses in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267–277.

Muenzinger, K. F. Concerning the effect of shock for right responses in visual discrimination learning. *J. exp. Psychol.*, 1948, **38**, 201–204.

Muenzinger, K. F., Brown, W. O., Crow, W. J., & Powloski, R. F. Motivation in learning: XI. An analysis of electric shock for correct responses into its avoidance

and accelerating components. *J. exp. Psychol.*, 1952, **43**, 115–119.

MUENZINGER, K. F., & FLETCHER, F. M. Motivation in learning: VII. The effect of an enforced delay at point of choice in the visual discrimination habit. *J. comp. Psychol.*, 1937, **23**, 383–392.

MUENZINGER, K. F., & NEWCOMB, H. Motivation in learning: V. The relative effectiveness of jumping a gap and crossing an electric grid in a visual discrimination habit. *J. comp. Psychol.*, 1936, **21**, 95–104.

MUENZINGER, K. F., & POWLOSKI, R. F. Motivation in learning: X. Comparison of electric shock for correct turns in a corrective and non-corrective situation. *J. exp. Psychol.*, 1951, **42**, 118–124.

PAVLOV, I. P. *Conditioned reflexes*. London: Oxford Univer. Press, 1927.

PRINCE, A. I. Effect of punishment on visual discrimination learning. *J. exp. Psychol.*, 1956, **52**, 381–385.

WISCHNER, G. J. The effect of punishment on discrimination learning in a non-correction situation. *J. exp. Psychol.*, 1947, **37**, 271–284.

(Received November 20, 1958)

Journal of Experimental Psychology
Vol. 58, No. 4, 1959

SUPPLEMENTARY REPORT: EFFECT OF INTERPOLATED UCS TRIALS IN EYELID CONDITIONING WITHOUT A READY SIGNAL[1]

K. P. GOODRICH, L. E. ROSS, AND A. R. WAGNER[2]

State University of Iowa

Kimble, Mann, and Dufort (1955) originally reported that UCS-alone trials interpolated in the middle of a series of CS-UCS conditioning trials produced a performance increment in eyelid conditioning comparable to that produced by paired conditioning trials. Goodrich, Ross, and Wagner (1957), however, were unable to obtain these results, and it was suggested that one of the factors differentiating the two experiments was the use of a ready signal by Goodrich, Ross, and Wagner and no ready signal by Kimble, Mann, and Dufort. Dufort and Kimble (1958) subsequently reported that, consistent with the original two studies, UCS-alone trials apparently led to a performance increment when no ready signal was employed, but failed to produce such an effect with a ready signal.

A pilot study by the present authors failed to show this effect of the no ready signal condition, the results being essentially like those found by Goodrich, Ross, and Wagner with a ready signal. It was therefore decided to conduct the present experiment in an attempt to provide information relevant to the determination of the relationship between the no ready signal condition and the effects of interpolated puff-alone trials in eyelid conditioning.

Method.—Eighty-nine women from an introductory psychology course served as *Ss*. Five of these were discarded as voluntary responders (Spence & Taylor, 1951). The apparatus, recording equipment, and procedure were essentially like those of the previous study, except that (*a*) the CS in the present experiment was the increase in luminance of a 2.25-in. milk glass disk from less than .004 to .32 apparent ft.-c. (instead of from .05 to 1.51 apparent ft.-c.), and (*b*) no ready signal was employed in the present experiment, the trials occurring at intervals of 15, 20, or 25 sec., according to a predetermined random order.

The *Ss* were assigned to three treatment groups with an attempt to match the frequency distributions of number of CRs on Trials 1–10. The *Ss* in Group C, the control group, received 60 successive paired conditioning trials. Group X-20 was identical to the experimental group in the previous study, receiving 20 conditioning trials, then 20 puff-alone trials, then 20 conditioning trials. Group X-10 was treated in the same manner as Group X-20, except that the 20 puff-alone trials were interpolated following Trial 10, followed by 30 conditioning trials. As in the original study, no instructions were given when the stimulus conditions were changed.

Results and Discussion.—Performance curves for the three groups are shown in Fig. 1. It may be seen that Groups X-10 and X-20 exhibit a lower level of performance following their interpolated UCS-alone trials than does Group C following a like number of conditioning trials. The distributions of scores were not normal, so nonparametric tests were used. The Kruskal-Wallis (1952) H test was significant ($H = 8.24$, $P < .02$) for the three groups over Trials 41–60. Mann-Whitney U tests between Group C and each of Groups X-10 and X-20 were significant ($P = .05$), but the difference between Groups X-10 and X-20 was not significant ($P > .20$). It seems clear that

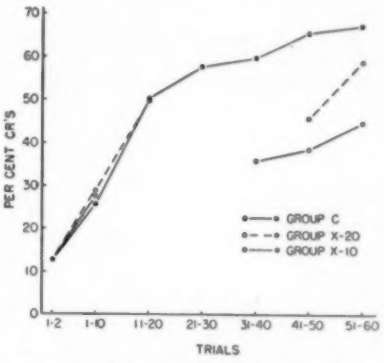


FIG. 1. Performance curves for the three groups ($N = 28$ per group).

[1] This study, carried out under the general direction of K. W. Spence, is part of a project concerned with the influence of motivation on performance in learning, under Contract N9 onr-93802, Project NR 154-107 between the State University of Iowa and the Office of Naval Research. During the course of the investigation, authors Goodrich and Wagner were National Science Foundation Predoctoral Fellows.

[2] Now at the University of Pennsylvania, the University of Wisconsin, and Yale University, respectively.

the puff-alone trials for the experimental groups did not have the same effect upon subsequent performance as did the same number of conditioning trials in the case of Group C.

On the reasoning that it is paired conditioning trials which are critical in performance increments, it is of interest to compare each of the two experimental groups with the control group at points in training at which they have had equal numbers of paired trials. Thus the performance of Group C on Trials 21–40 was compared with that of Group X-20 on Trials 41–60, and the performance of Group C on Trials 11–40 was compared with that of Group X-10 on Trials 31–60. It will be noted that the interpolated UCS-alone trials appeared for both experimental groups to produce a small decrement in performance, quite the opposite from the Dufort and Kimble results, although the differences were not statistically reliable ($P > .05$ in both cases). Similar results to the present ones were obtained in the original study with the counterparts of Groups C and X-20.

Dufort and Kimble reported that under conditions of no ready signal, as in the present experiment, the interpolated puff-alone trials resulted in subsequent performance comparable to that of a control group which had received all paired trials. The present results do not show this effect for either of two independent interpolated-UCS groups.

References

Dufort, R. H., & Kimble, G. A. Ready signals and the effect of interpolated UCS presentations in eyelid conditioning. *J. exp. Psychol.*, 1958, **56**, 1–7.

Goodrich, K. P., Ross, L. E., & Wagner, A. R. Performance in eyelid conditioning following interpolated presentations of the UCS. *J. exp. Psychol.*, 1957, **53**, 214–217.

Kimble, G. A., Mann, L. I., & Dufort, R. H. Classical and instrumental eyelid conditioning. *J. exp. Psychol.*, 1955, **49**, 407–417.

Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. statis. Ass.*, 1952, **47**, 583–621.

Spence, K. W., & Taylor, J. A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

(Received May 21, 1959)

LOOKING FOR A BOOK PUBLISHER?

It is no secret that publication is regarded as the foundation stone of a scholar's career. In order to gain recognition, prestige, and advancement, the scholar must seek publication.

Exposition Press offers scholars a complete publishing service, under our special academic imprint, *Exposition-University Books*. Though the proportion of scholarly publications in the lists of the trade houses is shrinking, due to economic pressures, several of our recent titles in the academic fields have gained popular reviews, professional recognition and steady sales. These books are the result of a publishing plan that is bringing a steady flow of scholarly writers to Exposition Press. They reflect the high editorial standards and quality of design and production which have won the respect of libraries, schools, booksellers and critics.

FREE! The behind-the-scenes story of book publishing revealed in two fact-filled, illustrated brochures, containing a detailed description of our 40%-royalty subsidy plan including a breakdown of contract terms and typical costs (in print for the first time). Copies are available on request. Your inquiries and manuscripts are invited. An editorial appraisal will be furnished promptly without obligation.

Please write to Jarold Paine, Editorial Department.

EXPOSITION PRESS, *Main Office, 386 Park Ave. S., N.Y. 16, N.Y.*